

QY	189	GGCGTCTCCGCGGCTCCAGGAGCCCGAGCTTGGCTGGCTTGCCTGCGCGCGCTGGTGAGC	248
Db	181	GGCGTCTCCGCGCTCCAGGAGCCCGAGCTTGGCTGGCTTGCCTGCGCGCGCTGGTGAGC	240
QY	249	ACTCGGCGGCGTGCAGCATGACCCCTGTGGAAACGGCGCTACTGCGCTTTTAAACCCCAAGCC	308
Db	241	ACTCGGCGGCGTGCAGCATGACCCCTGTGGAAACGGCGCTACTGCGCTTTTAAACCCCAAGCC	300
QY	309	CGGATATCCCGAGGCTTCAAGGTTCCACTGCTCATGCTATATCTGATATCTGAGTGTTTTGGCTCTA	368
Db	301	CGGATATCCCGAGGCTTCAAGGTTCCACTGCTCATGCTATATCTGATATCTGAGTGTTTTGGCTCTA	360
QY	369	GCAGCAAGCTTCCGCTCATCTTTCGCGGGATCCGTGGACACTGCGCGCTGTGGTTG	428
Db	361	GCAGCAAGCTTCCGCTCATCTTTCGCGGGATCCGTGGACACTGCGCGCTGTGGTTG	420
QY	429	GTCGAGGTTCTTCTCACTCTGTTCATAGCGCGAGAAATTGGCTGTGCACTTTCAGTGC	488
Db	421	GTCGAGGTTCTTCTCACTCTGTTCATAGCGCGAGAAATTGGCTGTGCACTTTCAGTGC	480
QY	489	GAAATGTTTCGGGGTTACAGTGAACCAACAACATCTTACAAAGGCTTCAGCGAGCGCG	548
Db	481	GAAATGTTTCGGGGTTACAGTGAACCAACAACATCTTACAAAGGCTTCAGCGAGCGCG	540
QY	549	GTTACAGCGCGCTGTCCGTTCTGCTGTGGGCGCTGGAGGGCATTAATATTAACCTCACAGG	608
Db	541	GTTACAGCGCGCTGTCCGTTCTGCTGTGGGCGCTGGAGGGCATTAATATTAACCTCACAGG	600
QY	609	ACCCCAATGATCAAGCTGAAAGAAACAATTAACATCAACAAGAGCATTCACTCGGCGCTG	668
Db	601	ACCCCAATGATCAAGCTGAAAGAAACAATTAACATCAACAAGAGCATTCACTCGGCGCTG	660
QY	669	AAAGAGATTACGCGCGCGAGTACGCGAACGCACTGAGAAAGGGCTGCCGACCCAGT	728
Db	661	AAAGAGATTACGCGCGCGAGTACGCGAACGCACTGAGAAAGGGCTGCCGACCCAGT	720
QY	729	CTCTACCTGGCGGAGAAAGTTCAACACGATGAGCCCTTGCAGGCTGTACACACATACCAC	788
Db	721	CTCTACCTGGCGGAGAAAGTTCAACACGATGAGCCCTTGCAGGCTGTACACACATACCAC	780
QY	789	CTGGCGGGACACTACGCTCGGACACGCTATAGGGTGGGCTTCTGCTCTCTCC	848
Db	781	CTGGCGGGACACTACGCTCGGACACGCTATAGGGTGGGCTTCTGCTCTCTCC	840
QY	849	AACGTGCTCTCTCCACGCGCGGCGCTTACGAGGCGCTGGACATGCTGACACCGGA	908
Db	841	AACGTGCTCTCTCCACGCGCGGCGCTTACGAGGCGCTGGACATGCTGACACCGGA	900
QY	909	GGCTTTCGGGCTCTTTCGGGGTCTTTCGGCTTTCGGCTTTCAGTGGGCGCTTCGCGG	968
Db	901	GGCTTTCGGGCTCTTTCGGGGTCTTTCGGCTTTCGGCTTTCAGTGGGCGCTTCGCGG	960
QY	969	CTCGGCTTAGGCTCTCTCGCGCTCACTAGTACGCGCGCGCTTTCGGGGTCAAGCTG	1028
Db	961	CTCGGCTTAGGCTCTCTCGCGCTCACTAGTACGCGCGCGCTTTCGGGGTCAAGCTG	1020
QY	1029	GCAACCGGCTCTCTGCTCTTCTCTGAGAGGGCGCTGGTGAATTCACATAGTTCCG	1088
Db	1021	GCAACCGGCTCTCTGCTCTTCTCTGAGAGGGCGCTGGTGAATTCACATAGTTCCG	1080
QY	1089	CCCAAGCGCTTCGACACCTTTCGACCAAAAGCGCAAGCATGCGACAGAGAGAGG	1148
Db	1081	CCCAAGCGCTTCGACACCTTTCGACCAAAAGCGCAAGCATGCGACAGAGAGAGG	1140
QY	1149	GGCTCACTTATCTTCGCGGACCCCATGCAACAAGCGCGCTCTTCCAGACTTAAAA	1208
Db	1141	GGCTCACTTATCTTCGCGGACCCCATGCAACAAGCGCGCTCTTCCAGACTTAAAA	1200
QY	1209	TGTATACCACTAACCTGAGAGGGGAGCCCAATCTGGACTCCTTCCCGGCTTGGAGCAT	1268
Db	1201	TGTATACCACTAACCTGAGAGGGGAGCCCAATCTGGACTCCTTCCCGGCTTGGAGCAT	1260

Qy	1269	CGAGAGCGCGGAAAGAGTGC	CCGCGACAGCGCTGGGCGCCAGAGAGAGTCCAGAGAAAGGCACTG	1328
Db	1261	CGAGAGCGCGGAAAGACATGTCC	CGCCAGGCGCTGGCGCCAGAGAGAGCTCCAGAGAAAGGCACTG	1320
Qy	1329	AGCGCTGCTGGCGCGAGAGCGCT	CGGACATCCGACAGGACCCAGGAAAGTCTTCTCTGGGCGGA	1388
Db	1321	AGCGCTGCTGGCGCGAGAGCGCT	CGGACATCCGACAGGACCCAGGAAAGTCTTCTCTGGGCGGA	1380
Qy	1389	TCGTGAATTAACCTTTT	TTCTTTTCTTTTCTTTT	1421
Db	1381	TCGTGAATTAACCTTTT	TTCTTTTCTTTTCTTTT	1413
RESULT 2				
LOCUS	HSN806241	1740 bp	mRNA	linear PRI 17-JUN-2003
DEFINITION	Homo sapiens mRNA; cDNA DKFP686C04213 (from clone DKFP686C04213).			
ACCESSION	BX37581			
VERSION	BX537581.1 GI:31873659			
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Rukavota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 1740)			
TITLE	Bloecker,H., Boecker,M., Mewes,H.W., Well,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and Wiemann,S.			
JOURNAL	Direct Submission			
COMMENT	Submitted (17-JUN-2003) MRP5, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY			
FEATURES	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBR (National Research Centre for Biotechnology Ltd., Brunsheweig/Germany) within the cDNA sequencing consortium of the German Genome Project.			
source	This clone (DKFP686C04213) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/. location/Qualifiers			
1..1740	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/map="15q15"			
	/clone="DKFP686C04213"			
	/tissue_type="human rectum tumor"			
	/clone_lib="686 (synonym: hlccs). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"			
	/dev_stage="adult"			
	1694..1699			
	1720			
polyA_signal	polyA_site			
ORIGIN	71.1%; Score 1010.6; DB 8; Length 1740;			
Query Match	Best Local Similarity 98.6%; Pred. No.1-le-166;			
Matches 1019; Conservative	0; Mismatches 14; Indels 0; Gaps 0;			
Db	688	CTGACCCGGGTCATTCCTG	CAGCGCTGGTTTGGTGGAGAGTCTCTCAGTCT	747
Qy	449	GTTTCATAGGCGCAGAAATTGTG	GTGCTGTGCATTCAGTGCAGAGAAATGTTCTGGGATACGT	508
Db	748	GTTTCATAGGCGCAGAAATTGTG	GTGCTGTGCATTCAGTGCAGAGAAATGTTCTGGGATACGT	807
Qy	509	GAAACCAACACATCTCA	CAAGCTTCAGGCGACGCGGCTTACAGCCCGTGTGAGTCT	568
Db	808	GAAACCAACACATCTCA	CAAGCTTCAGGCGACGCGGCTTACAGCCCGTGTGAGTCT	867
Qy	569	GCTCTGTGGGCTTGAGGCGCATTA	TATTACACTACAGGAAACCCAGTGACATCAGCTGA	628

```

Db      868 GCTCGTGGGCTGAGGAGCATTAATTAACCTACAGAGGAGCCCACTGATCAGCTGAA 927
Qy      629 CGAGACATTGACTACACAGAGATTACCTTGGCTGTAAAGAGATTACCCCGGGA 688
Db      928 CGAGACATTGACTACACAGAGATTACCTTGGCTGTAAAGAGATTACCCCGGGA 987
Qy      689 GTACCGGAAAGCACTGAGAAAGGGGCTGCGGAGCCGAGTCTTACCTGGCGGAGAGTT 748
Db      988 GTACCGGAAAGCACTGAGAAAGGGGCTGCGGAGCCGAGTCTTACCTGGCGGAGAGTT 1047
Qy      749 CACACGAGTAGGCTTGGGCTGTACACAGTACCACTTGGCGGAGACATACGCTC 808
Db      1048 CACACGAGTAGGCTTGGGCTGTACACAGTACCACTTGGCGGAGACATACGCTC 1107
Qy      809 GGCACGCTATGGGTGGCTTGTGCTTCTGCTCTTCCACAGTGTCTTCCACGCTC 868
Db      1108 GGCACGCTATGGGTGGCTTGTGCTTCTGCTCTTCCACAGTGTCTTCCACGCTC 1167
Qy      869 GGCCTCGCTTACGAGAGCTTGGCACTGCACTGCACTGCGAGCCTTGGCGCTTCCGCGGT 928
Db      1168 GGCCTCGCTTACGAGAGCTTGGCACTGCACTGCACTGCGAGCCTTGGCGCTTCCGCGGT 1227
Qy      929 CTTGCGCTTGGGCTTCACTCTCTAGCGTGCCTGCGCTGCGCTGAGGCTTCCGCG 988
Db      1228 CTTGCGCTTGGGCTTCACTCTCTAGCGTGCCTGCGCTGCGCTGAGGCTTCCGCG 1287
Qy      989 GCTTACCACTCAAGTACGCGCGCTTCTGAGTCACTGCGCACTGCGCTTCTGCT 1048
Db      1288 GCTTACCACTCAAGTACGCGCGCTTCTGAGTCACTGCGCACTGCGCTTCTGCT 1347
Qy      1049 CTTTCTTGGAGGGGCTTGTGTGATCTTCCAGTATGTTCCGCGCTTCTTCCAGCT 1108
Db      1348 CTTTCTTGGAGGGGCTTGTGTGATCTTCCAGTATGTTCCGCGCTTCTTCCAGCT 1407
Qy      1109 TCTGGAACAAGCGCAAGAGCTGACAGCAAGAGAGAGGGGCTCACTCTTATCTCTCG 1168
Db      1408 TCTGGAACAAGCGCAAGAGCTGACAGCAAGAGAGAGGGGCTCACTCTTATCTCTCG 1467
Qy      1169 CGACCACTGCACAAGAGCGCGCTTCCAGCTTAAATGATATCACTTAACTGTG 1228
Db      1468 CGACCACTGCACAAGAGCGCGCTTCCAGCTTAAATGATATCACTTAACTGTG 1527
Qy      1229 AGGGGGAACCAATTTGATCTTCCCGCTTGGGACATGCGAGGCGCGGAGACATGTC 1288
Db      1528 AGGGGGAACCAATTTGATCTTCCCGCTTGGGACATGCGAGGCGCGGAGACATGTC 1587
Qy      1289 CCGGCAAGGCTGGGCAAGAGCTCCAGAGAAAGGCACTGAGCGTGTGCGGAGAGCC 1348
Db      1588 CCGGCAAGGCTGGGCAAGAGCTCCAGAGAAAGGCACTGAGCGTGTGCGGAGAGCC 1647
Qy      1349 TCGGACATCCGACAGGCACTGAGAAAGTCTCTGGGCGAGTCTGTAATAAACTTTT 1408
Db      1648 TCGGACATCCGACAGGCACTGAGAAAGTCTCTGGGCGAGTCTGTAATAAACTTTT 1707
Qy      1409 TCTTTTGTTTT 1421
Db      1708 TCTTTTGTTTT 1720

```

```

RESULT 3
LOCUS      CQ727948          741 bp    DNA          linear    PAT 03-FEB-2004
DEFINITION Sequence 13882 from Patent W002068579.
ACCESSION  CQ727948
VERSION     CQ727948.1  GI:42295393
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homidae; Homo.
REFERENCE   1

```

```

AUTHORS    Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE      K1ts, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL    Patent: WO 02068579-A 13882 06-SEP-2002;
            PB Corporation (NY) (US)
FEATURES   Location/Qualifiers
            source          1..741
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN
Query Match      52.1%; Score 741; DB 6; Length 741;
Best Local Similarity 100.0%; Pred. No. 1.6e-119;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      485 TGCAGATGGTTGTTGGGTACAGTGAACCAACATCTTACAAAGCTTCAAGCGAGC 544
Db      1 TGCAGATGGTTGTTGGGTACAGTGAACCAACATCTTACAAAGCTTCAAGCGAGC 60
Qy      545 GCGCGTTACAGCCCGTGTGCTTGTCTGCGGCGCTGAGGGGCAATTAATTAACACTAC 604
Db      61 GCGCGTTACAGCCCGTGTGCTTGTCTGCGGCGCTGAGGGGCAATTAATTAACACTAC 120
Qy      605 AGGGAACCCAGTGCATCAGCTGAACGAGACATTTGACTACAAAGAGCACTTCACTGCG 664
Db      121 AGGGAACCCAGTGCATCAGCTGAACGAGACATTTGACTACAAAGAGCACTTCACTGCG 180
Qy      665 TCTGAAGAAGATTACGCGCGGAGTACGCGAACGCACTGAGAAAGGGCTGCGGAGCC 724
Db      181 TCTGAAGAAGATTACGCGCGGAGTACGCGAACGCACTGAGAAAGGGCTGCGGAGCC 240
Qy      725 AGTGTCTTACTGCGGAGAAATTCAACCGAGTAGGCTTGGGCTGTATACACAGTA 784
Db      241 AGTGTCTTACTGCGGAGAAATTCAACCGAGTAGGCTTGGGCTGTATACACAGTA 300
Qy      785 CCACTGGGAGGACACTACGCTCGGCACTATGAGGTGAGCTTGTGCTTGGCTCT 844
Db      301 CCACTGGGAGGACACTACGCTCGGCACTATGAGGTGAGCTTGTGCTTGGCTCT 360
Qy      845 CTCGAAGTGTCTCTTCAAGCGCGCGCTTCAAGAGGCTGAGCACTGTGACAC 904
Db      361 CTCGAAGTGTCTCTTCAAGCGCGCGCTTCAAGAGGCTGAGCACTGTGACAC 420
Qy      905 CGAGCTTGTGCGCTTCTGCGGGTCTTGGCTTGGCTTCAATCTTATGCTGCGCTGTG 964
Db      421 CGAGCTTGTGCGCTTCTGCGGGTCTTGGCTTGGCTTCAATCTTATGCTGCGCTGTG 480
Qy      965 CCGGCTCGGCTTGGGCTCTCCGCGCTCACTACATGATGAGGCGCGCTTGGGTCAAC 1024
Db      481 CCGGCTCGGCTTGGGCTCTCCGCGCTCACTACATGATGAGGCGCGCTTGGGTCAAC 540
Qy      1025 GCTGGAACAAGCGGTCTGTGCTCTTCTGAGAGGGGCGTGTGAGTCTCCAGTAGT 1084
Db      541 GCTGGAACAAGCGGTCTGTGCTCTTCTGAGAGGGGCGTGTGAGTCTCCAGTAGT 600
Qy      1085 TCGGCGCAAGGCTTGTGCACTTCTGAGACCAAGGCGCAAGGACTGACGCGAGAG 1144
Db      601 TCGGCGCAAGGCTTGTGCACTTCTGAGACCAAGGCGCAAGGACTGACGCGAGAG 660
Qy      1145 AGGGGGCTCACTTTATCTCGGCGACCACTGCAACAAGAGGCGCTCTCCAGACTT 1204
Db      661 AGGGGGCTCACTTTATCTCGGCGACCACTGCAACAAGAGGCGCTCTCCAGACTT 720
Qy      1205 AAATGTATCACTAAGCT 1225
Db      721 AAATGTATCACTAAGCT 741

```

```

RESULT 4
LOCUS      BC031111          1354 bp    mRNA          linear    ROD 06-JUL-2005

```

DEFINITION Mus musculus RIKEN cDNA 9030623N16 gene, mRNA (cDNA clone MGC:35715 IMAGE:4989815), complete cds.

ACCESSION BC031111

VERSION BC031111.1 GI:21411407

KEYWORDS MGC

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1354)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klauer R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buicow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang D., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Rahe S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulyahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez J., Helton E., Kettelman M., Madan A., Rodriguez S., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E., Scherck A., Schein J.E., Jones S.J. and Marra M.A.
Mammalian Gene Collection Program Team
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1354)

CONSTRM TITLE NIH MGC Project

JOURNAL Direct Submission

PUBMED Submitted (03-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

AUTHORS NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

CONSTRM COMMENT Contact: MGC help desk
Email: gcabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdcpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

REMARK Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 58 Row: 9 Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13365245.

FEATURES

source Location/Qualifiers

1..1354

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:35715 IMAGE:4989815"

/issue_type="Colon, normal, 5 month old male mouse."

/clone_id="NCI CGAP_Co24"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1354

/gene="9030623N16R1k"

/db_xref="GeneID:66811"

/db_xref="MGI:1914061"

CDS

129..1091

/gene="9030623N16R1k"

/codon_start=1

/product="Numb-interacting protein 2"

/protein_id="AAH31111.1"

/db_xref="GI:21411408"

/db_xref="GeneID:66811"

/db_xref="MGI:1914061"

/translation="MTAMDVLPYPPYPPRHAASFVPLIYLIVFLSLAFLFLIPG
IRGSRFMVLPVPLSLSTGAIYAVHSGPMFVGRVMTNTSYAFSPRQVYGLH
VGLAGVNLITLGTPEQDLNFTIDINERTFMNEDYTRKYHALKEGIPDPVLYAEK
FTPSGPNCLYQHYHLAGYAAATLVACFWIINAALSMPLVPLGLGAVLITGFTL
FGVFAPASISVPLCHFRLAGAVLPYVGASFMTLATGILSLGLGAVLIDHYRPS
ALRSLDLSVDCSNQAKNSPLTINPQHQLKSPDLINTLL"

ORIGIN

Query Match	48.7%	Score 691.4	DB 9	Length 1354
Best Local Similarity	75.8%	Pred. No. 6.9e-111		
Matches 913	Conservative	0	Mismatches 276	Indels 16; Gaps 4
230	CTGCCCGCCTGGCTGACAGACTCGGCGGCGTGCAGCATGACCTGTGGACGGGCTACT			289
92	CTCGCTGCTGGAAACAGTTCGAACTTTCGAGATGATCGGTGGAGCGGGTGTCT			151
290	GCCCTTTTACCCCAAGCCCGGCGATGCGGAGGCTTTCAGCTTCACTGCTCATGCTTAT			349
152	ACCCCTTTTACCCCAAGCCCGGCGATGCGGAGGCTTTCAGCTTCACTGCTCATGCTTAT			211
350	TCTAGTGTGTTTGGCTCTAGCAGCAAGCTTCTGCTCATCTTGGCGGGGATCGGGCCA			409
212	CTGGGTGTTCTGTCTTGGCTGCTAGCTTCTGCTCATCTTGGCGGGGATCGGGCCA			271
410	CTCGCGCTGTTTGGTGTGAGAGTCTTCTCACTGCTTTCATAGCGGCAAAATGT			469
272	CTCGCGCTGTTCTGTTGTGTAAGTCTTCTCACTGCTTTCATAGCGGCAAAATGT			331
470	GAGCTGCACTTCAAGTGCAAAATGTTCTGGTACAGTGAACCAACATCTCTACA			529
332	GAGCTGCACTTCAAGTGCAAAATGTTCTGGTACAGTGAACCAACATCTCTACA			391
530	AGCTTCAAGCGAGCGGGTTCACGCCGTGCGTCTGCTGGGCGCTGAGGGCAT			589
392	AGCTTCAAGCGAGCGGGTTCACGCCGTGCGTCTGCTGGGCGCTGAGGGCAT			451
590	TATATTACATCAACGAGACCCCAAGTGCATGATGAAAGAGACCATGATCAACGA			649
452	TATATTACATCAACGAGACCCCAAGTGCATGATGAAAGAGACCATGATCAACGA			511
650	GCAATTCACCTGGCTGTGAAGAAATTAAGCGCGGAGTACGGAACGCACTGGAGA			709
512	GCGTTTCACTTGGCTGTGAAGAAATTAAGCGCGGAGTACGGAACGCACTGGAGA			571
710	GGGCTGCGGGAACCAAGTGTCTACCTGGCGGAGAGTTTCAACGAGTACCTTGGCG			769
572	GGGCTGCGGGAACCAAGTGTCTACCTGGCGGAGAGTTTCAACGAGTACCTTGGCG			631
632	GCTGACCAATATCACTGCGCGGTCACTATGCGCGCAACTGCGGTGCAATT			691
830	CTGCTTGTGCTCTCTCCCAAGTGTCTCTCAAGCGCGCGCGCTCTACGAGAGGCT			889
692	CTGCTTGTGCTCTCTCTCCCAAGTGTCTCTCAAGCGCGCGCGCTCTACGAGAGGCT			751
890	GGAAGCTGACACACCGAGCTTGGCGCTTCCGAGGCTTTCGCTTGGCTCTCATCTC			949
752	GGAAGCTGACACACCGAGCTTGGCGCTTCCGAGGCTTTCGCTTGGCTCTCATCTC			811
950	TAGGTGCGCTCTGCGCGCTTCCGCTTACGCTCTCTCCGCTTCACTAGTACGCGC			1009
812	CAGGTGCGCTCTGCGCGCTTCCGCTTACGCTCTCTCCGCTTCACTAGTACGCGC			871
1010	GCGCTTGTGAGTACGCTGGAACCGCGCTCTGTGCTCTTCTCTGGAGCGGCGGTGT			1069

```

Db      872 CTCCTTTGGCTACCGTGGCCACCGGCATCTGAGCTCTCTCGGAGGGCGGTGT 931
Qy      1070 GAGCTTCAGATATGTTGCGCCAGCGCTCTTGCAACCTTTGAGCAACCAAGCCCAAGGA 1129
Db      932 GATTCTCACTATATACCGCCGACCGCCTTCGCTCTTCTGAGATCTAAGTGTCAAGA 991
Qy      1130 CT---GAGCGCAGAGAGAGAGGGGCTCACTCTTATCTCTGGGAGCCCAATGCAAGCA 1186
Db      992 CTGTACGACACAGAGCTAAGAGAACTCACTCTCAACCTCAACCAACCCGCAACAGCA 1051
Qy      1187 GCGCGCTCTCCAGACTTAATATGATCAACCACTAATCTGAGAGGGGAGCCCAATGCA 1246
Db      1052 GTTGAAGAGCTCCAGACTTA---TATTACCACTCTCTGTGAAAATACTGACTCCGA 1108
Qy      1247 CTCCTTCCCGCCTTGGAAGATGCGACGCGGGAAGAGTGCCTGAGGCTTG-GGCCA 1305
Db      1109 TTTCTACCCCTCTTTGGAGCCCATAGACCTGAAGACTGTATTAGAGCGCGTGCAGGA 1168
Qy      1306 GAGAGCTCCAGAGAGGCGACTGAGCGCTGTCGCGGAGCGCTCGACATCCGCAAGCA 1365
Db      1169 GCATAGCTGAGCTGAGAACTGCTGTCTCCCTGCGGAGCAAGAGGGGCAACCCCATGTG 1228
Qy      1366 CC-----AGGGAAGCTCTCGGAGGAGATCTGTAATAAATCTTTTCTTTGT 1416
Db      1229 CTTTGTCTAGAGAAATGAGATTTCTCAGAGAACTGTAAATAATCTTTTGTCTTTT 1288
Qy      1417 TTTT 1421
Db      1289 TTTT 1293

RESULT 5
AF547266 5851 bp mRNA linear MAM 03-Apr-2003
LOCUS     Sus scrofa dual oxidase 1 (Duox1) mRNA, complete cds.
DEFINITION
ACCESSION AF547266
VERSION    AF547266.1 GI:23664372
KEYWORDS
SOURCE
ORGANISM  Sus scrofa (pig)
          Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
          Sus.
REFERENCE 1 (bases 1 to 5851)
AUTHORS   Morand,S., Chazaoui,M., Kaniewski,J., Deme,D., Ohayon,R.,
          Noel-Hudson,M.S., Virion,A. and Dupuy,C.
TITLE      Effect of iodide on nicotinamide adenine dinucleotide phosphate
          oxidase activity and Duox2 protein expression in isolated porcine
          thyroid follicles
JOURNAL    Endocrinology 144 (4), 1241-1248 (2003)
PUBMED     12639906
REFERENCE 2 (bases 1 to 5851)
AUTHORS   Kaniewski,J., Morand,S., Noel-Hudson,M.-S., Ohayon,R., Virion,A.
          and Dupuy,C.
TITLE      Direct Submission
JOURNAL    Submitted (13-SEP-2002) Unite 486, INSERM University of Paris-Sud,
          5 rue JB Clement, Chateau-Malabry 92296, France
FEATURES
SOURCE     location/Qualifiers
           1..5851
           /organism="Sus scrofa"
           /mol_type="mRNA"
           /db_xref="taxon:9823"
           1..5851
           /gene="Duox1"
           948..5609
           /gene="Duox1"
           /EC_number="1.6.3"
           /function="putatively generates H2O2 by transferring
           electrons from NADH/NADPH to molecular oxygen"
           /function="oxidoreductase acting on NAD(P)H with oxygen as
           acceptor"
           /note="NADPH:O2 oxidoreductase"

```

```

/codon start=1
/product="dual oxidase 1"
/protein_id="F03933.1"
/db_xref="GI:23664373"
/translation="MGRFLALMTLLVGRMMKPAKNSISWVQRPDGMNIMMKY
GSKGRLQILVPASVADGVYQIPAGEPLINPDLSTNARKPGAGLSLRVTVGVF
GYHVLSDLVISIEKPGPAEFLNIHIPGDPVDPNKSQVVLPPQRNDPNTQSPS
NPRDLINENYGLMDSAIYSSHSMDLRSPGQGLASGAPDPAFPRQDPLPMPT
PDPAQGRGPDGLYAFAGQGRNRPFLMLCAQILMFYHNLCAQILREHLMQDDEL
QHARKVITATYQSIITMYELPSTIRKPMQRYGVRPLDPSISPEPLASBQFSTV
PQVYMRNASCHFOGINRNSVSRLARCNSTNSRSHENRPLQAEVDLLLMASOI
ASEREDMVEDVDQFWPGLKPSRTDLASCLQRGDLIPSTTKARAGLPVTRW
ODINPALSRSRSGIVLEATLALYNQDLRLPGLLESJGDPLESTIVLQFVRL
RQDGRWFENFTKNGLFSEKIEIRMTSIRDVLAATNMPTALQNVFEMHAGDPC
QPROSLTKPLACAPLIMDYPGSGRFGVGTGLTCEPLVLLLSAMLYAOLRRNF
KRLQVONROSIMCKLVGMCKLEMGRCRQPCPVVHLOSGQIHMWDGLSVLTQI
LRPQOVNLLLSNHRKTLTLKIPRYDLVLMFPLDEBERQVENVLQSLKSGSP
OEWELEQELMRALVREORSHLETFFHPLFQVLDIDQADGALPLDSQVREML
TCELSRAPFASLGLKPDMPVMSFSLADKQNGYLSFRPLDILVFNKGSBERKS
RLMFMYDQDNGLSKDFIRMLRSPFIEISNLSKAOLEVESEMFRAQODKOR
LTMEDFHMRLRQDSELPTOLCVKGVVEPVTKDLCRASYSIOEKLCSPPVSAHC
PNSNDVBYELTWMKIQCPDTPDPOEIRRRPKRYTSPQLLPTAHREKFSRSH
QTVQPKRFEVENTRRHIGLAVFTYIAGLFLERATYVPAAHMGIOTTRGIIIS
RGTAAISFMFSYILLTMCRNLITFLRETFLNKXVPPDAVDRILASTAIIILTVLH
SAGVVNVVLYLSISPLVSLCPGLFHDNGSEFPQKYWMFPQVGLVWMLLIL
ALMYVPSAHHFRCSFRGFWLTHLYLILYMLLIIGSPALIQLPFHILFVLPAIY
VGRKLYSARKYVESVVKAEILLPSGVTHLOPORPGFYEKSGQWVRIACLALGTEY
HPFTLSAPREDTSLHIRAGPMTTRHEIYSPPINDCNKPKYLXQDPRSGHQR
MHRFVSIVLVGGIGVTPASILKDLVFNSSVSCQVCKIYITWRTIROGBWLD
IIRVEVNDHRDLVSVAIYITQAEKFDLATTMLYICERFQVLYNRSLFTGLASLTH
FGRPPPEFPFNSIQEVAHPQVRKIGVSCSPGPMTKXVEKACOLINRQDTRFISHYEN
F"

ORIGIN
Query Match 44.1%; Score 626.2; DB 4; Length 5851;
Best Local Similarity 81.8%; Pred. No. 1,4e-99;
Matches 735; Conservative 0; Mismatches 163; Indels 1; Gaps 1;

262 GCAGCATGACCTCTGTGGAAGGGGTACTGCTTTTATCCCGCCGCGGAGATCCGCGAG 321
Db 10 GCAGCATGACCTGTGTGGAAGGGGTGCTGCTTTTATCCCGCCGCGGAGATCCGCGCA 69
Qy 322 GCTTCAGCGTTTCACTGCTCATGCTTATCTAGTGTGTTTGGCTTCTGAGCAGCACTTCC 381
Db 70 GCTTCAGCGTTTCACTGCTCATGCTTATCTGAGTGTGTTTGGCTTCTGAGCAGCACTTCC 129
Qy 382 TGCTCATCTTGCGCGGGATCGTGCGCACTCGCGCTGTTTGGTGTGTGAGACTTTC 441
Db 130 TACTCATCTTGCGCGGGATCGTGCGCACTCGCGCTGTTTGGTGTGTGAGACTTTC 189
Qy 442 TCACTCTGTTCTATGCGCGCAAAATTTGGTCTGCGACTTCACTGCGAAGATGTTGCG 501
Db 190 TCACTCTGTTCTATGCGCGCAAAATTTGGTCTGCGACTTCACTGCGAAGATGTTGCG 249
Qy 502 GTACAGTGAACACCAACACATCTTCAAAAGCTTCAAGCGGCGCGGCTTCAAGCCCGT 561
Db 250 GCGAGTTGAGCAGAAATCAATCTTCAAGGCTTCAAGTGTGCGACGTTGCCAACCCACC 309
Qy 622 AGCTGAACGAGACCATTTGATCAACAGAGAGTTCACTGCGCTGTGAAGAGAAATTACG 681
Db 310 TCGGTCTGCAAGTGGCGCTGAGAGGCTTAATATTACACTCAAGGAAATCCGTTACAGC 369
Qy 682 TCGGTCTGCTGTGGCGCTGAGAGGCTTAATATTACACTCAAGGAAATCCGTTACAGC 429
Db 370 AGCTGAATGAGACCATTTGATCAACAGAGAGTTCACTGCGCTGTGAAGAGAAATTACG 489
Qy 742 AGAGTTTCAACCGAGTATGCTTGGCGCTTGAACCAAGTATCACTGCGCGGAGACT 801
Db 490 AGAGTTT-ACCTCCAGAAAGCCCTGCGGGGTTTACCCGCAATATGCTTACCGGAGACT 548

```

QY	802	ACGCCCTGGGCAAGCAGTAAGGAGGCGCTTCGTGCTTCGGCTCTCTTCTTCAACGTCGCTCT	861
Db	549	AGCCTCCGCGCAGATTGTGAGACGGCGCTTCGTCTTGACTCTCTTCCAACTGCTGCTTT	608
QY	862	CCAGCGCGGCGCCCGCTCTAAGAGAGGCTGGCACTGTGACCAACCGAGGCTTCGCGCTCT	921
Db	609	CCATCCCGGTCCCGGATTAAGAGGCGCTGCTCTCTCTCATCAACGGCGGCTTCGCGCTCT	668
QY	922	TGCGGGTCTTCGCTTCGCGCTCATCTCTAAGAGTGGCGGCTTCGCGCGCTCCGCGCTAAGCT	981
Db	669	TCTGTGTGTGGCTTCGCTCTTACTCCAGGTGCTCTTTCGCAAGCTCCGCGCTCGGCT	728
QY	982	CCTCGGCGCTCAACCACTCAGTACGGGCGCGCTTCGTGGTCAAGCTGGCAACCGGCGCTTC	1041
Db	729	CTTCGGAAGTCAACCACTCACTAAGGTGCTGCTTTTGGATCAAGCTGGCGCACCGGCGATTC	788
QY	1042	TGTGCTTCTCTCGAGAGGGGCGGTGTGATGCTCCAGATATTTCGGCCGACGCTCTTC	1100
Db	789	TGTGCTCTCTCCTCGAGAGGCGGTGTGATGATCTTCACTAGCGCTGACCCGAGCTCTTC	848
QY	1102	GCAACCTCTGAGACCAAGCGCGCAAGACTGAGCCAGAGAGAGGGGCTTCACCTCTT	1160
Db	849	GCACTTCTTAAAGGAGATCAAGAGCTGGATATGTCAAGGCAGCAAAAGGCTCTCCATTT	907

RESULT 6	AX035348	498 bp	DNA	Linear	PAT 15-NOV-2000
LOCUS	AX035348				
DEFINITION	Sequence 3 from Patent WO0053748.				
ACCESSION	AX035348				
VERSION	AX035348.1	GI:11191065			
KEYWORDS					
SOURCE	Homo sapiens (human)				

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.	Novel compounds	Patent : WO 0053748-A 3 14-SEP-2000;
	BRUCK CLAUUDINE ELIVRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE)		VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART JEAN POL (BE)
FEATURES	source	Location/Qualifiers	
		1..498	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	

	Query Match	31.9%;	Score 453;	DB 6;	Length 498;
	Best Local Similarity	99.6%;	Prod. No. 3.6e-65;		
	Matches 475;	Conservative 0;	Mismatches 0;	Indels 2;	Gaps 2;
QY	947	CTTAAAGCTGCGCGCTCTGCCCGCT-CCGCTTAAGGCTCTTCGCGCTCAACACTCAATGAC	1005		
Db	1	CTTAAAGGTCCCGCTCTGCCCGCTCCCGCTTAAGGCTCTTCGCGCTCAACACTCAATGAC	60		
QY	1006	-GGCGCGCTTCTTGAGTCAAGCTGACAGCGAGCGGCGCTTGAGCGCTTCTTGAGAGGGCC	1064		
Db	61	AGCGCGCGCTTCTTGAGTCAAGCTGACAGCGAGCGGCGCTTGAGCGCTTCTTGAGAGGGCC	120		
QY	1065	GTGGTAGTCTCAAGTATGTTGGGCCAGGCGCTTTGACACCCCTTGTGACCAAGCGCC	1124		
Db	121	GTGGTAGTCTCAAGTATGTTGGGCCAGGCGCTTTGACACCCCTTGTGACCAAGCGCC	180		
QY	1125	AAGGACTGACGCCAGGAGAGAGGGGGCTCACTTTATCTTGAGGACCCACATGCACAAG	1184		
Db	181	AAGGACTGACGCCAGGAGAGAGGGGGCTCACTTTATCTTGAGGACCCACATGCACAAG	240		
QY	1185	CAGGGCGCTCTCCCAACTTAAATGATATCAACACTGATGAGGGGAGCCCAATCTTG	1244		

Db	241	CAGGCGCGCTCCGAGACTTAAATATGATACCACTAACCTGTGAGGGGAGCCCAATCTG	300
Qy	1245	GACTCCTTCCCCGGCTTGGGACATTCGACAGGCGGGGAAAGCAATGGCCCGCAGGCTTGGGCC	1304
Db	301	GACTCCTTCCCCGGCTTGGGACATTCGACAGGCGGGGAAAGCAATGGCCCGCAGGCTTGGGCC	360
Qy	1305	AGGAGAGCTCAGGAAAGGCACTGAGGCGTCTGGCGAGGCTTCGACATTCGCAAGGC	1364
Db	361	AGGAGAGCTCAGGAAAGGCACTGAGGCGTCTGGCGAGGCTTCGACATTCGCAAGGC	420
Qy	1365	ACGAGGGAAGTCTCCTGGGGCGATCTGTAAATAACCTTTTTCCTTGTGTTTTT	1421
Db	421	ACGAGGGAAGTCTCCTGGGGCGATCTGTAAATAACCTTTTTCCTTGTGTTTTT	477

RESULT 7	AC012255	166937 bp	DNA	1linear	HTG 06-MAY-2001
LOCUS	AC012255				
DEFINITION	Homo sapiens chromosome 15 clone Rpl1-109p20 map 15, WORKING DRAFT SEQUENCE, 22 unordered pieces.				
ACCESSION	AC012255				
VERSION	AC012255.4	GI:7684440			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 166937)	Blairten, B., Linton, L., Nuebaum, C. and Lander, E.	Homo sapiens chromosome 15, clone RP11-109D20	Unpublished
2 (bases 1 to 166937)			
3	Blairten, B., Linton, L., Nuebaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boughalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Domingo, M., Doneelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Healdorf, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karats, A., Klein, J., Lowndez, J., Llew, C., Locke, K., MacDonald, P., Margulis, N., McEwen, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, C. N., Subramanian, A., Talmac, J., Tesfaye, S., Titrell, A., Vassiliev, H., Vo, A., Wheeler, T., Wu, X., Wymen, D., Ye, W., Zimmer, A. and Zody, M.	Direct Submission	
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

```

http://ftp.genome.washington.edu/BM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: 13737
Center clone name: 109_D_20
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 153800 bases at least Q40
Consensus quality: 160606 bases at least Q30
Consensus quality: 163375 bases at least Q20
Insert size: 164837; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
2087: contig of 2087 bp in length
2088
2187: gap of 100 bp
2188
3626: contig of 1439 bp in length
3627
3726: gap of 100 bp
3727
4845: contig of 1119 bp in length
4846
4945: gap of 100 bp
4946
6702: contig of 1757 bp in length
6703
8326: contig of 1524 bp in length
8327
8426: gap of 100 bp
8427
10494: contig of 2068 bp in length
10495
10594: gap of 100 bp
10595
13215: contig of 2621 bp in length
13216
13315: gap of 100 bp
13316
16451: contig of 3136 bp in length
16452
16551: gap of 100 bp
16552
20843: contig of 4292 bp in length
20844
20943: gap of 100 bp
20944
24940: contig of 3997 bp in length
24941
25040: gap of 100 bp
25041
29901: contig of 4861 bp in length
29902
30001: gap of 100 bp
30002
34189: contig of 4188 bp in length
34190
34289: gap of 100 bp
34290
38386: contig of 4097 bp in length
38387
38486: gap of 100 bp
38487
45827: contig of 7341 bp in length
45828
45927: gap of 100 bp
45928
51729: contig of 5802 bp in length
51730
51829: gap of 100 bp
51830
60123: contig of 8294 bp in length
60124
60223: gap of 100 bp
60224
73532: contig of 13309 bp in length
73533
73632: gap of 100 bp
73633
87433: contig of 13801 bp in length
87434
87533: gap of 100 bp
87534
101708: contig of 14175 bp in length
101709
101808: gap of 100 bp
101809
121193: contig of 19385 bp in length
121194
121293: gap of 100 bp
121294
139051: contig of 17758 bp in length
139052
139151: gap of 100 bp
139152
166937: contig of 27786 bp in length.
Location/Qualifiers

FEATURES
source
1. 166937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-109D20"
/clone_id="RPC1-11 Human Male BAC"

misc_feature
1. 2087
/note="assembly_fragment"
gap
2088
/estimated_length=100
misc_feature
2188
/estimated_length=100
gap
3627
/estimated_length=100
misc_feature
3727
/estimated_length=100
gap
4846
/estimated_length=100
misc_feature
4946
/estimated_length=100

gap
/note="assembly_fragment"
6703
/estimated_length=100
misc_feature
6803
/note="assembly_fragment"
8327
/estimated_length=100
gap
8426
/estimated_length=100
misc_feature
8427
/note="assembly_fragment"
10495
/estimated_length=100
gap
10594
/estimated_length=100
misc_feature
10595
/note="assembly_fragment"
13215
/estimated_length=100
gap
13216
/estimated_length=100
misc_feature
13316
/note="assembly_fragment"
16451
/estimated_length=100
gap
16452
/estimated_length=100
misc_feature
16552
/note="assembly_fragment"
20843
/estimated_length=100
gap
20844
/note="assembly_fragment"
20943
/estimated_length=100
misc_feature
20944
/note="assembly_fragment"
24940
/estimated_length=100
gap
24941
/estimated_length=100
misc_feature
25041
/note="assembly_fragment"
29902
/estimated_length=100
gap
30002
/estimated_length=100
misc_feature
34190
/note="assembly_fragment"
34289
/estimated_length=100
gap
34290
/estimated_length=100
misc_feature
38387
/note="assembly_fragment"
38486
/estimated_length=100
gap
38487
/note="assembly_fragment"
45827
/estimated_length=100
misc_feature
45828
/note="assembly_fragment"
45927
/estimated_length=100
gap
45928
/estimated_length=100
misc_feature
51730
/note="assembly_fragment"
51829
/estimated_length=100
gap
51830
/estimated_length=100
misc_feature
60123
/note="assembly_fragment"
60223
/estimated_length=100
gap
60224
/estimated_length=100
misc_feature
73533
/note="assembly_fragment"
73632
/estimated_length=100
gap
73633
/note="assembly_fragment"
87433
/estimated_length=100
misc_feature
87434
/note="assembly_fragment"
87533
/estimated_length=100
gap
87534
/estimated_length=100
misc_feature
101709
/note="assembly_fragment"
101808
/estimated_length=100
gap
101809
/note="assembly_fragment"
121193
/estimated_length=100
misc_feature
121194
/note="assembly_fragment"
121293
/estimated_length=100
gap
121294
/note="assembly_fragment"
139051
/estimated_length=100
misc_feature
139052
/note="assembly_fragment"
139151
/estimated_length=100
gap
139152
/estimated_length=100

Query Match	29.1%	Score 413.4	DB 14	Length 166937
Best Local Similarity	98.6%	Pred. No. 1,36-62		
Matches 417	Conservative	0	Mismatches 6	Indels 0
			Gaps	0
QY 1	GAACGAGGAAAGTAA	CGGCTACAGACAGT	GAGAAATAGTTTGGCTTCGCGCGGCTAGAAAA	60
Db 157708	GAACGAGGAAAGTAA	CGGCTACAGACAGT	GAGAAATAGTTTGGCTTCGCGCGGCTAGAAAA	157767
QY 61	CTGTGCGGTACCA	CCCCAGAGCGTTGAGAGACGCCACCTCCAGCGTTCTTACGGA	120	
Db 157768	CTGTGCGGTACCA	CCCCAGAGCGTTGAGAGACGCCACCTCCAGCGTTCTTACGGA	157827	
QY 121	GAGGTGACGAGACT	CAGACTTTCACACAGCCCACTCGTCCAGCGTTGTACGCAAAAGACG	180	
Db 157828	GAGGTGACGAGACT	CAGACTTTCACACAGCCCACTCGTCCAGCGTTGTACGCAAAAGACG	157887	
QY 181	CCAGAGAGCGCGCT	CTCCCGGCTTCAGAGAGCCCCAGACTTGTGGCTTGGCGCGCGCTG	240	
Db 157888	CCAGAGAGCGCGCT	CTCCCGGCTTCAGAGAGCCCCAGACTTGTGGCTTGGCGCGCGCTG	157947	
QY 241	CGTGACGAGACTCG	CGCGCGGTGACAGATACCTGTGGAACGGCGTATCGCTTTTACC	300	
Db 157948	CGTGACGAGACTCG	CGCGCGGTGACAGATACCTGTGGAACGGCGTATCGCTTTTACC	158007	
QY 301	CCGACCCCCGCGAT	CGCGAGCTTTCAGGTTTCACTGCTCATCGTATTTCTAGTGT	360	
Db 158008	CCGACCCCCGCGAT	CGCGAGCTTTCAGGTTTCACTGCTCATCGTATTTCTAGTGT	158067	
QY 361	TGGCTCTAGCAGAC	AGCTTCTGCTCATCTTTCGCGGAGATCGTGGCACTCCGCGCTGT	420	
Db 158068	TGGCTCTAGCAGAC	AGCTTCTGCTCATCTTTCGCGGAGATCGTGGCACTCCGCGCTGT	158127	
QY 421	TTT 423			
Db 158128	TGT 158130			
RESULT 8	AC091117	181312 bp	DNA linear	PRI 20-FEB-2002
LOCUS	AC091117	181312 bp	DNA linear	PRI 20-FEB-2002
DEFINITION	Human chromosome 15 clone RP11-109D20 map 15q15, complete sequence.			
ACCESSION	AC091117			
VERSION	AC091117.5	GI:18767395		
KEYWORDS	HTG.			
SOURCE	Human sapiens (human)			
ORGANISM	Human sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 181312)			
AUTHORS	Rosen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D., and Hood, L.			
TITLE	Sequencing of human chromosome 15 D1S146-D1S117 region			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 181312)			
AUTHORS	Rosen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D., and Hood, L.			
TITLE	Direct SubMISSION			
JOURNAL	Submitted (30-MAR-2001) Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA			
REFERENCE	3 (bases 1 to 181312)			
AUTHORS	Rosen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D., and Hood, L.			
TITLE	Direct SubMISSION			
JOURNAL	Submitted (13-FEB-2002) Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA			
REFERENCE	4 (bases 1 to 181312)			

```

AUTHORS      Rowen,L., Madan,A., Qin,S.,Baradarani,L., Birditt,B., Bloom,S.,
              Burke,J., Dora,M., Fleetwood,P., Kaur,A., Madan,A., Neabitt,R.,
              Pate,D. and Hood,L.
TITLE        Direct Submision
JOURNAL      Submitted (14-FEB-2002) Multimegabase Sequencing Center, Institute
              for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
              98105, USA
REFERENCE    5 (bases 1 to 181312)
AUTHORS      Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
              Burke,J., Dora,M., Fleetwood,P., Kaur,A., Madan,A., Neabitt,R.,
              Pate,D. and Hood,L.
TITLE        Direct Submision
JOURNAL      Submitted (20-FEB-2002) Multimegabase Sequencing Center, Institute
              for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
              98105, USA
COMMENT      On Feb 20, 2002 this sequence version replaced gl:18653547.
              ----- Genome Center
              Center: Multimegabase Sequencing Center
              Center code: UMSC
              Web site: http://chroma.mbt.washington.edu/meg_www
              Contact: leerowen@systemsbiology.org
              ----- Summary Statistics
              Sequencing vector: pUC18; 108752
              Chemistry: Dye-terminator Big Dye; 90% of reads
              Chemistry: Dye-primer Big Dye; 10% of reads
              Assembly program: Phrap; version 0.990399

Note: Data from overlapping clone AC090888 [Drafting center: WIBR]
was added for finishing. RP11-109D20 spans bases 1-174236 of this
sequence. Bases 174237-179372 derive from clone RP11-276K9,
AC087790 [Drafting center: WIBR] to establish the overlap with the
next BAC in our tiling path. Bases 179373-181312 derive from
finished clone CTD-2651B20, AC051619 [Drafting center: UMSC] in
order to give overlap required for determining long range
contiguity.
FEATURES             Location/Qualifiers
     source            1..181312
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="15"
                        /map="15q15"
                        /clone="RP11-109D20"
                        /clone_id="RP11-109D20"
                        /note="This clone overlaps CTD-2014N11 AC090888 and
                        RP11-276K9 AC087790. Data from overlapping BACs were added
                        and the consensus sequence determined from RP11-109D20 to
                        the extent possible."
     misc_feature      1..89137
                        /note="overlap with CTD-2014N11, AC090888"
     unsure            36107..36120
                        /note="low quality data"
     misc_feature      40470..40820
                        /note="sequence data generated from subcloned PCR product"
     misc_feature      53030..53480
                        /note="sequence data generated from subcloned PCR product"
     unsure            86982..86983
                        /note="low quality data"
     misc_feature      95610..96030
                        /note="sequence data generated from subcloned PCR product"
     misc_feature      96830..97420
                        /note="sequence data generated from subcloned PCR product"
     misc_feature      98260..98420
                        /note="sequence data generated from subcloned PCR product"
     unsure            /note="single clone coverage"
     unsure            98850..98852
                        /note="low quality data"
     unsure            100981..101003
                        /note="low quality data"
     misc_feature      174237..179372
                        /note="overlap with RP11-276K9, AC087790. This sequence
                        data was added to establish the overlap between the BACs
                        in the tiling path"

```


misc_feature 179373..181312
/note="overlap with CTD-2651B20, AC051619. This data is added to give overlap for long range contiguity."

ORIGIN

Query Match 29.1%; Score 413.4; DB 8; Length 181312;
Best Local Similarity 98.6%; Pred. No. 1.2e-62;
Matches 417; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

TITLE
JOURNAL

COMMENT

Testaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, W.
Submitted (28-Aug-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gt:6056272.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2336

Center clone name: 163.P.10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 108837 bases at least Q40

Consensus quality: 145227 bases at least Q20

Insert size: 157000; agarose-1p

Insert size: 155134; sum-of-coverage

Quality coverage: 3.2 in Q20 bases; agarose-1p

Quality coverage: 3.3 in Q20 bases; sum-of-coverage

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1170	1169: contig of 1169 bp in length
*	1269	gap of 100 bp
*	1270	3335: contig of 2066 bp in length
*	3336	3435: gap of 100 bp
*	3436	5698: contig of 2263 bp in length
*	5699	5798: gap of 100 bp
*	5799	9216: contig of 3418 bp in length
*	9217	9316: gap of 100 bp
*	9317	15633: contig of 6317 bp in length
*	15634	15733: gap of 100 bp
*	15734	21746: contig of 6013 bp in length
*	21747	21846: gap of 100 bp
*	21847	30351: contig of 8505 bp in length
*	30352	30451: gap of 100 bp
*	30452	35509: contig of 6058 bp in length
*	35510	36609: gap of 100 bp
*	36610	45279: contig of 8670 bp in length
*	45280	45379: gap of 100 bp
*	45380	56851: contig of 11472 bp in length
*	56852	56951: gap of 100 bp
*	56952	67374: contig of 10423 bp in length
*	67375	67474: gap of 100 bp
*	67475	79534: contig of 12060 bp in length
*	79535	79634: gap of 100 bp
*	79635	95432: contig of 15798 bp in length
*	95433	95532: gap of 100 bp
*	95533	120801: contig of 25269 bp in length
*	120802	120901: gap of 100 bp
*	120902	156534: contig of 35633 bp in length.

FEATURES

source

1..156534
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"

RESULT 9
AC009700/c 156534 bp DNA linear HTG 20-APR-2000
LOCUS Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
DEFINITION
AC009700
ACCESSION AC009700.4 GI:7622346
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckwith, R., Breen, J., Brown, A.,
Cacace, A., Cerny, J., Colangelo, M., Collins, S., Collins, A.,
Cooke, P., DeRubeis, K., Depayre, E., Devos, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardina, S., Gilbert, D., Grant, G.,
Hagood, B., Heath, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
Karunas, A., Lebeck, J., Liu, C., Locke, K., Macdonald, P.,
Maratias, A., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meislin, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,

QY 361 TGCGCTTACGACGACCTTCTGCTCATCTTGGCGGGGATCCGTGGSCCACTGCGCTGCT 420
 DB 140765 TGCGCTTACGACGACCTTCTGCTCATCTTGGCGGGGATCCGTGGSCCACTGCGTAAAGG 140824
 QY 421 TTT 423
 DB 140825 TGT 140827

RESULT 11
 AX714511/c 2684 bp DNA linear PAT 15-APR-2003
 LOCUS
 DEFINITION Sequence 1195 from Patent EPI293569.
 ACCESSION AX714511
 VERSION AX714511.1 GI:29889464
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE
 AUTHORS Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechiki, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.
 TITLE Full-length cDNAs
 JOURNAL Patent: EP 1293569-A 1195 19-MAR-2003;
 Helix Research Institute (Jp) ; Research Association for Biotechnology (Jp)
 FEATURES
 source location/Qualifiers
 1..2684
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 27.5%; Score 391; DB 6; Length 2684;
 Best Local Similarity 97.5%; Pred. No. 1,9e-58;
 Matches 397; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1015 TCTGGGTCAAGCTGCAACCGGCGCTCTGTGCTCTTCTTGGAGGGCGCGTGTGATC 1074
 DB 2263 TTTTGATCCCAACCGGCGGCGCTCTGTGCTCTTCTTGGAGGGCGCGTGTGATC 2304
 QY 1075 TCCAGTATGTTGGCGCCGCGCTCTTGGCAACCTTTCGACCAAGGCGCAAGGACTGCA 1134
 DB 2303 TCCAGTATGTTGGCGCCGCGCTCTTGGCAACCTTTCGACCAAGGCGCAAGGACTGCA 2244
 QY 1135 GCCAGGAGAGAGGGGGCTCACTCTTATCTCTGGCGACCACTGACAGAGGCGCGCTC 1194
 DB 2243 GCCAGGAGAGAGGGGGCTCACTCTTATCTCTGGCGACCACTGACAGAGGCGCGCTC 2184
 QY 1195 TCCGAGACTTAATGATATCACTAACTGTGTAGGGGGACCAATTTGATCTCTTCC 1254
 DB 2183 TCCGAGACTTAATGATATCACTAACTGTGTAGGGGGACCAATTTGATCTCTTCC 2124
 QY 1255 CCGGCTTGGGACATTCGAGGCGCGGAGACATGCGCGGCGCGGCGCGGAGAGCTC 1314
 DB 2123 CCGGCTTGGGACATTCGAGGCGCGGAGACATGCGCGGCGCGGCGCGGAGAGCTC 2064
 QY 1315 CAGGAAGGGGCACTGAGCGCTGTGGCGGAGGCGCTCGACATTCGACAGGCAACAGGAAA 1374
 DB 2063 CAGGAAGGGGCACTGAGCGCTGTGGCGGAGGCGCTCGACATTCGACAGGCAACAGGAAA 2004
 QY 1375 GTCTCTGGGGCGATCTGTAAATTAACCTTTTCTTTTGTGTTTTT 1421
 DB 2003 GTCTCTGGGGCGATCTGTAAATTAACCTTTTCTTTTGTGTTTTT 1957

RESULT 12
 AK056896/c

LOCUS AK056896 2684 bp mRNA linear PRI 30-JAN-2004
 DEFINITION Homo sapiens CDNA FLJ32334 file, clone PROST2005426.
 ACCESSION AK056896
 VERSION AK056896.1 GI:16552419
 KEYWORDS oligo capping, f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE

AUTHORS

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

CDS

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PROST2005426"
/issue_type="prostate"
/clone_id="PROST2"
/notes="cloning vector: pMR18SFL3"
56..952
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB71304.1"
/db_xref="GI:16552420"
/translation="MATTGHTPEFVAGPKPTPEMDTTLASIMIFLATATPVLIVPG
IGKTRFLPMLRVVSTLPIGAMILGPPVOOLNHTIYNSEFTRGENTAEKAKLSE
KCLPPVYLAKEKPPSPGCGLYROYRLGHGYSAMLMWAPICMLANMLAMPVYV
GGMLATITPOLALLPFSMATSTSPCPHIGASVLTTHSPAWITLTGSLCIVL
LGIAMVAVHMQPRLKAFPNOSVDSDPMLSESPERGALLSPKTRMSADSPYSQDIPL
SEASTKAYCKEAKHPDPDCL"

```

ORIGIN

```

Query Match      27.5%; Score 391; DB 8; Length 2684;
Best Local Similarity 97.5%; Pred. No. 1.9e-58;
Matches 397; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1015 TCTGGTCAAGCTGGCAACCGGCTGCTGCTTCTTCTCGAGGAGGCGGTGAGTC 1074
DB 2363 TTTTCATCCCAACCCGCAAGGCTGCTGCTTCTTCTCGAGGAGGCGGTGAGTC 2304
QY 1075 TCCAGTATGTTGGCCCAAGGCTCTTTCGACCTTTCGACCAAGGCCAGAGCTGCA 1134
DB 2303 TCCAGTATGTTGGCCCAAGGCTCTTTCGACCTTTCGACCAAGGCCAGAGCTGCA 2244
QY 1135 GCGAGAGAGAGGGGCTCACTCTTATCTGCGGCAACCGCAAGGAGGCGGCTC 1194
DB 2243 GCGAGAGAGAGGGGCTCACTCTTATCTGCGGCAACCGCAAGGAGGCGGCTC 2184
QY 1195 TCCAGACTTAAATGATATCACTAACCTGAGGGGAGCCCAATCTGACTCTCTTC 1254
DB 2183 TCCAGACTTAAATGATATCACTAACCTGAGGGGAGCCCAATCTGACTCTCTTC 2124
QY 1255 CCGCTTGGAGATCGCAAGGCGGAGAGAGGTCGCCGCGCTGCGGCGAGAGAGCTC 1314
DB 2123 CCGCTTGGAGATCGCAAGGCGGAGAGAGGTCGCCGCGCTGCGGCGAGAGAGCTC 2064
QY 1315 CAGGAAGGCGATCGAGCGCTGCTGCGGCGGAGGCTCGGACATCCGAGGAGCAAGGAAA 1374
DB 2063 CAGGAAGGCGATCGAGCGCTGCTGCGGCGGAGGCTCGGACATCCGAGGAGCAAGGAAA 2004
QY 1375 GTCTCCGGGGGCGATCTGTAATAAAGCTTTTCTTTTCTTTTCTTTTCTTTT 1421
DB 2003 GTCTCCGGGGGCGATCTGTAATAAAGCTTTTCTTTTCTTTTCTTTTCTTTT 1957

```

RESULT 13

AC009700

LOCUS Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
DEFINITION Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
AC009700 156534 bp DNA linear HTG 20-APR-2000
SEQUENCE 15 unordered pieces.

ACCESSION

AC009700

VERSION AC009700.4 GI:7622346
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.

SOURCE

Homo sapiens

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 156534)
Birten, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-163P10
Unpublished
2 (bases 1 to 156534)
Birten, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Becker, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,

TITLE
JOURNAL
COMMENT

Cooke, P., Dearellano, K., Depayre, B., Devon, K., Dewar, K.,
Donlan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funk, R., Gage, D., Galagan, J., Gardys, S., Gilbert, D., Grant, G.,
Haggs, B., Heath, A., Horton, L., Howard, J. C., Jones, C., Kam, L.,
Karlsson, A., Lechowicz, J., Liu, C., Locke, K., Macdonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nioff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Testa, S., Torrella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, W.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 163 P 10

Center clone name: 163 P 10

Sequencing Vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 108837 bases at least Q40

Consensus quality: 130230 bases at least Q30

Consensus quality: 145227 bases at least Q20

Insert size: 157000; agarose-fp

Insert size: 155134; sum-of-contigs

Quality coverage: 3.2 in Q20 bases; agarose-fp

Quality coverage: 3.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1169:	contig of 1169 bp	in length
1170	1269:	gap of 100 bp	
1270	3335:	contig of 2066 bp	in length
3336	3435:	gap of 100 bp	
3436	5698:	contig of 2263 bp	in length
5699	5798:	gap of 100 bp	
5799	9216:	contig of 3418 bp	in length
9217	9316:	gap of 100 bp	
9317	15633:	contig of 6317 bp	in length
15634	15733:	gap of 100 bp	
15734	21746:	contig of 6013 bp	in length
21747	21846:	gap of 100 bp	
30351	30351:	contig of 8505 bp	in length
30352	30451:	gap of 100 bp	
30452	36509:	contig of 6058 bp	in length
36510	36609:	gap of 100 bp	
36610	45279:	contig of 8670 bp	in length
45280	45379:	gap of 100 bp	
45380	56851:	contig of 11472 bp	in length
56852	56951:	gap of 100 bp	
56952	67374:	contig of 10423 bp	in length
67375	67474:	gap of 100 bp	
67475	79534:	contig of 12060 bp	in length
79535	79634:	gap of 100 bp	
79635	95432:	contig of 15798 bp	in length
95433	95532:	gap of 100 bp	

Matches	556;	Conservative	0;	Mismatches	311;	Indels	3;	Gaps	1;
QY	263	CAGCATGACCTTGTGGAACGGGCTACTGGCTTTTACCCCGAGCCCGGCGATGCCGAG	322						
DB	437	CAGAGTGGTACTTGTGGGACACATTCCTCTTATGTGGGCCCAACCACTTCCC	496						
QY	323	CTTCAGCGTTTCCACTGCTATCGTTATTTAGTGTGGCTCTGAGCAAGTTCCT	382						
DB	497	GATGACACCACTTTTGGCAGCATCATGATCTTTTGTGACGTGACACCGCTTCA	556						
QY	383	GCTCATCTTGGCGGAGTCCGTGCGCACTCGCGTGTGGTTTGGTGTGAGATTCCT	442						
DB	557	CGTCATCGGCGGATTCGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	616						
QY	443	CAGTCTGTTATAGGCGGAGAAATTTGGTGTGACCTTCACTGAGCAATGTTCGGG	502						
DB	617	CAGCTTATTCATCGGGGCTGCAATCTGTGGTGAATTTCAAGTTGAGGTGTGTGG	676						
QY	503	TACAGTGAACCAACACATTCCTCAAAAGGCTTCAAGGAGCGGCGTTACAGCCGTG	562						
DB	677	CCAGTCAAGACCAACATCATCAAGGCTTCAAGTTGAGTGAATCAGCGCTGAT	736						
QY	563	CGGTCTGCTGTGGGCTGTGAGGAGCAATTAATATTAACAACAAGCAAGCAAGCA	622						
DB	737	TGGGCTGCAAGTGGGCTGGGTGGATGACATCACTCAAGGAGACCCCGTGAGCA	796						
QY	623	GCTGAACGAGACCATTAATCAACAAGCAAGTTCACCTGGCGTCTGAAGAATTA	682						
DB	797	GCTGAATGAGACCATTAATCAACAAGCAAGTTCACCTGGCGCTGGTGAAGTAT	856						
QY	683	CGGAGATGACCAACGCACTGAGAGAGAGGCGCGAGACCAAGTGTCTACCTGGG	742						
DB	857	TAGAGATGAGCAAGAGCTCTGAGAGAGGCGCGAGACCAAGTGTCTACCTGA	916						
QY	743	GAGTTCACACGAGTACCTTGTGGGCTGTATCAACAAGTTCACCTGGCGAGCA	802						
DB	917	GAGTTCACCTCCAGAGAGGCGCTGTATACCGCAATGACCGCTGGCGGAGCA	976						
QY	803	CGCTTGGGCAAGCTATGGTGTGGTGTCTCTTCTGCTCTCTCAACGTCTGCTC	862						
DB	977	CACCTGAGCATGCTATGGTGTGGGCTCTCTGCTGCTCTGCGCAATGTGCTCT	1036						
QY	863	CAGCGCGCGCGCGCTCTACAGAGAGCGCGGAGCACTGAGCAACAAGGAGCTT	922						
DB	1037	CAGGCTGCGCGCTGTATATGCTGTGCTATGCTATGCTATGCTATGCTATG	1096						
QY	923	CG---GGGTCTTGGGCTTGGCTCTCATCTCTAGCGTGCCTGCGCTGCGCTAG	979						
DB	1097	GGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1156						
QY	980	CTCTCTGCGAGCTCAACAAGTACAGGCGCGCTTGTGGGTCAAGTGTGCAAC	1039						
DB	1157	CGCTTGTGCTGTATCTCAACATGAGGCGCTTCTGTGATCAATTAACACAG	1216						
QY	1040	CTGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1099						
DB	1217	GCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1276						
QY	1100	TGCGACCTTCTGTGAGCAAGAGCGGCAAGGA	1129						
DB	1277	GAAAGCTTCTTCAACAGAGTGTGATGA	1306						

RESULT 15
 BC029819
 LOCUS BC029819 1923 bp mRNA linear PRI 30-JUN-2004
 DEFINITION Homo sapiens homolog of Drosophila Numb-interacting protein, mRNA
 (CDNA clone MGC:35310 IMAGE:517619), complete cds.
 ACCESSION BC029819
 VERSION BC029819.1 GI:20987583
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 1923)
 Strassberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, M., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothyluki, S., Garinchi, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

2 (bases 1 to 1923)
 Strassberg, R.D.
 Direct Submission
 Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 NIH-MGC help desk
 Contact: cgabs-r@mail.nih.gov
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: angbcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

REMARK

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAC Plate: 51 Row: 0 Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389332.

FEATURES

1..1923
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35310 IMAGE:517619"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_lib="NIH MGC_115"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..1923

gene

/gene="NIP"
 /note="synonym: FLJ32334"
 /db_xref="GeneID:90527"
 401..1852

CDS

/gene="NIP"
 /codon_start=1
 /product="Numb-interacting protein"
 /protein_id="AAH29819.1"
 /db_xref="GI:20987584"

/db_xref="GeneID:90527"
/translation="MATLGHPPFVAGPKPTPMDTLASIIIMPLTALATRVILG
IRGKRLFWLIRVYVSLFTGAALIANVSGSVGVSTNTSKASSSMIGADIDLO
VGGGVNITITGTPVOILNETITNREFTRLQENYAEYAKLEGLDPIVYLLK
FTPRSPGLRQRLAGHYTSMWAFICWILANWYMPVLYGGYMLATGIFOL
LALFPMATISLSPCLHIGASVLIHTHGPAMWITLTLGLCVLGLMAVAHRQP
HRLKAFNOSVDEPMLWSPREBGLSPRYSMASPKSODPLSEASTKAYRPR
RLSLVPADVRSALPGALISALPGALIAOARLALPGLRCPAGKESRLGPPSPMRGP
EGGERMAHHTGDSPLRIGRGTRLMWGSKERACGRAMLPLVNSGLKRPSC
DLPKMDYRDADAFFHLEPTVTSRHTPL"

ORIGIN

Query Match 25.1%; Score 356.4; DB 8; Length 1923;
Best Local Similarity 63.9%; Pred. No. 2.2e-52;
Matches 556; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

QY 263 CAGCATGACCCCTGTGAAAGCGCTACTGCTTTTACCCCGAGCCGCGATGCCGAG 322
DB 397 CAAGATGGCTACTTTGGGACACATCCCTCTATGCTGGCCCAAGCCAACCTTCCC 456
QY 323 CTTGAGGTTCCACTGCTCATGTTATTTAGTGTTTGGCTCTAGCAGCAGCTTCC 382
DB 457 GATGACACCACTTTGGCCAGCATCATGATCTTTCGACTGCACTGGCCAGCTTCAT 516
QY 383 GCTCATCTTGCCGAGGATCCGTCGCCACTGCTGTTTGGTGTGAGATCTTCT 442
DB 517 CGTCATCTGCTGGCATTCGGGAAAGACGAGGCTGTTCTGGCTGCTCGGGTGTGAC 576
QY 443 CAGTCTGTCATAGGGGCAAAATTTGGCTGTGCACTTCAGTGCAGAAATGTTGGTGG 502
DB 577 CAGCTTATTCATGGGGCTGCAATCCGCTGTAAATTCAGTTCGAGTGTCTGTGG 636
QY 503 TACAGTGAACACCAACATCTTACAAAGCCTCAGCGCAGCGCGCTTACAGCCGCT 562
DB 637 CAGGTCAAGACCAACATCATCAAGGCTTCAGTTCTGAGTGATCAGGCTGATAT 696
QY 563 CGGCTGCTGTGGGCTTGGAGGGCATTAATATTAACACTCAGAGGACCCCACTGATCA 622
DB 697 TGGGCTGCAGAGTGGGCTGGGTGAGTCAACATCACTCAAGGACCCCGTGCAGCA 756
QY 623 GCTGAACGAGACCATTTGACTTACAAAGAGGCTTCACTGGCTGTGAAGAATTAACG 682
DB 757 GCTGAATGAGACCATTAATTAACAAGAGGCTTCACTGGCTGTGAGAACTATGC 816
QY 683 CGCGAGTACGCGAAGCACTGGAAGGGGCTGCCGACCCAGTCTCTAAGCTGCGGA 742
DB 817 TAGAGAGTATGCAAAAGGCTCTGGAAGAGGCTGCCAGACCTGTGTGACTGACTGA 876
QY 743 GAAATTCAACCGAGTACCTTTCGCGCTGTACCAACAGTACCACTTGGCGGACACTA 802
DB 877 GAAATTCACTCCAAAGGCCATGTGCGCTATACCGCCAGTACCGCTTGGCGGACACTA 936
QY 803 CGGCTGGGCAAGCTATAGGTGGGCTTCTGCTTCTGCTCTCCATCAAGTGTGCTTC 862
DB 937 CACTTAAGCCATGCTATAGGTGGCATTTCTCTGTGCTGCTGCGCCAAATGTATGCTTC 996
QY 863 CAGCGGCGCCCGCTCTACGAGGCTTGGCACTGTGACCAACCGAGCCTTGGCGCTT 922
DB 997 CAGGCTGTGCTGTATATGTGTGCTACATGTATTTGGCCAGGCACTTTCAGGCTGTT 1056
QY 923 CG--GAGTCTTGGCTTGGCTTCCATCTCTAGCGTGCCTGCTGCGCTTACG 979
DB 1057 GAGCTGTGCTTCTTCTTCCATGGCCACATCACTCAACCTGTCTCCCTGCACTGG 1116
QY 980 CTCCTCCGCGCTCAACACTAGTACGAGCGCGCTTCTGGGTCAAGCTGGCAACGCGGCT 1039
DB 1117 CGCTTGTGTGCTCATATCAATGAGCTTGTGATACATTAACACAGGACT 1176
QY 1040 CCTGTGCTCTTCTCTCGAGGGGCGGTGTGAGTCTCAGATATGTTCCGCGCCAGCGCTCT 1099
DB 1177 GCTGTGTGTGCTGTGGCTGTGCTATGGCGGTGGCCACAGGATGAGGCTCAGAGGCT 1236
QY 1100 TGGCACCTTCTTGGACCAAGCGCAAGGA 1129

DB 1237 GAAAGCTTCTTCAACCAAGCTGTGATCA 1266

Search completed: January 22, 2006, 00:49:52
Job time : 7522 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 19:32:01 ; Search time 937 Seconds
(without alignments)
10107.287 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaagtaacgctc.....ctttttcttttctttt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneseqn1990s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1421	100.0	1421	3	AAA96505
2	1421	100.0	1450	12	ADL06425
3	1418.6	99.8	1474	6	ABL90535
4	1413	99.4	1441	3	AAA94623
5	1412	99.4	1491	10	ADD19228
6	1375.8	96.8	1460	3	AAC95564
7	1375.8	96.8	1460	8	ABZ73640
8	1375.8	96.8	1460	10	ADC20289
9	1375.8	96.8	1460	10	ABZ67235
10	958.4	67.4	960	13	ADU01671
11	768.8	54.1	981	10	ACF58262
12	768.8	54.1	981	12	ADO85846
13	453	31.9	498	3	AAA94624
14	413.4	29.1	5033	8	ABZ74587
15	413.4	29.1	5033	10	ADC21005
16	413.4	29.1	5033	10	ABZ68115
17	413	28.3	580	10	ADD19187
18	402.8	27.5	406	5	AA56563
19	391	27.5	2684	10	ACF58261

C	20	391	27.5	2684	10	ADA53627	Human cod
C	21	391	27.5	2684	12	ADN06056	Antipsori
C	22	391	27.5	2684	12	ADO85844	Colon can
C	23	356.4	25.1	1636	6	ABN59923	Novel hum
C	24	354.8	25.0	1029	5	AA45134	Human INT
C	25	354.8	25.0	1029	8	ACD66787	Secreted
C	26	354.8	25.0	1029	12	ADQ10402	Human pol
C	27	354.8	25.0	1029	12	ADQ98274	Human pol
C	28	354.8	25.0	2133	5	AA45133	Human INT
C	29	354.8	25.0	2133	8	ABX94107	cDNA enco
C	30	354.8	25.0	2133	8	ACD66786	Secreted
C	31	354.8	25.0	2133	12	ADQ10401	Human pol
C	32	354.8	25.0	2133	12	ADQ98273	Human pol
C	33	348	24.5	348	10	ACF58260	AA1821606
C	34	348	24.5	348	12	ADO85843	Colon can
C	35	299.4	21.1	2439	10	ADG32833	Human DNA
C	36	245.8	17.3	2684	10	ACF58261	ENST00000
C	37	245.8	17.3	2684	10	ADA53627	Human cod
C	38	245.8	17.3	2684	12	ADN06056	Antipsori
C	39	245.8	17.3	2684	12	ADO85844	Colon can
C	40	161.8	11.4	1338	5	AA567707	DNA enco
C	41	65.6	4.6	1602	4	ABL21213	Drosophi1
C	42	65.6	4.6	1808	4	ABL07775	Drosophi1
C	43	64	4.5	512	4	AAH36185	Human col
C	44	60.6	4.3	2000	8	ADA71938	Rice gene
C	45	52.6	3.7	134	8	ABZ74588	Secreted

ALIGNMENTS

RESULT 1	AAA96505	standard; cDNA, 1421 bp.
ID	AAA96505	standard; cDNA, 1421 bp.
XX	AAA96505;	
AC	08-FEB-2001	(first entry)
XX		
DT		
XX		
DE		
XX		
XX		
KW	Human; transmembrane protein; cell proliferation disorder; myeloma;	
KW	reproductive disorder; smooth muscle disorder; neurological disorder;	
KW	arteriosclerosis; leukaemia; acquired immunodeficiency syndrome; AIDS;	
KW	allergy; ovulatory defect; angina; hypertension; stroke; epilepsy.	
KW	Alzheimer's disease; Tourette's disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
PH		
FT	Key	Location/Qualifiers
FT	CDS	267..1229
XX		/*tag= a
PN		
XX		
PD	WO200056891-A2.	
XX		
XX	28-SEP-2000.	
XX		
XX	22-MAR-2000; 2000WO-US007817.	
XX		
XX	22-MAR-1999; 99US-0125537P.	
PR	16-JUN-1999; 99US-0139565P.	
XX		
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Yue H, Lal P, Tang YT, Hillman JL, Reddy R, Bandman O;	
XX	Baughn WR, Lu DM, Azimzal Y, Yang J;	
XX		
DR	WPI: 2000-579485/54.	
XX	P-PSDB; AAB18992.	
XX		
PT	New human transmembrane proteins are used to treat a disease or condition	
PT	associated with decreased expression of functional HTPP e-9. Tourette's	
PT	disorder, angina and leukemia.	

XX Claim 4; Page 129; 130pp; English.

CC The present sequence encodes a human transmembrane proteins (HMP).
 CC Agonists and antagonists of the protein are used to treat a disease or
 CC condition associated with overexpression of the protein. Diseases and
 CC conditions which can be treated include cell proliferative,
 CC immunological, reproductive, smooth muscle and neurological disorders
 CC e.g. arteriosclerosis, myeloma, leukemia, acquired immunodeficiency
 CC syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,
 CC stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The
 CC polynucleotides may be used to detect and quantify gene expression in
 CC biopsied tissues where protein expression may be correlated with disease
 CC e.g. to determine absence, presence or excess expression of HMP or to
 CC monitor regulation of HMP expression during therapeutic intervention

XX Sequence 1421 BP; 264 A; 468 C; 395 G; 294 T; 0 U; 0 Other;

Query Match 100.0%; Score 1421; DB 3; Length 1421;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACGAGAAAGTACGCTACACAGACAGTGAATAATGTTGCTGCGCGCTAGAAAA 60
 Db 1 GAACGAGAAAGTAAAGGCTACAGACAGTGAATAATGTTGCTGCGCGCTAGAAAA 60
 QY 61 CTCTGTGGTACCAACCCGAGAGCTTGAAGAGCGCCACTCCAGGCTTCTTAACGA 120
 Db 61 CTCTGTGGTACCAACCCGAGAGCTTGAAGAGCGCCACTCCAGGCTTCTTAACGA 120
 QY 121 GAGGTGACAGACTCAGACTTACCAAGCCACTGCTCCAGCCTTGTAGCAAGAGAG 180
 Db 121 GAGGTGACAGACTCAGACTTACCAAGCCACTGCTCCAGCCTTGTAGCAAGAGAG 180
 QY 181 CCAAGGACCGGCTCTCCCGGTGACAGAGCCCACTGCTGCTGCTGCTGCTGCTG 240
 Db 181 CCAAGGACCGGCTCTCCCGGTGACAGAGCCCACTGCTGCTGCTGCTGCTGCTG 240
 QY 241 CGTGACGACTCGGCGGCGGTGACAGATGACCTGTGGAACGGCGGTACTGCTTTTAC 300
 Db 241 CGTGACGACTCGGCGGCGGTGACAGATGACCTGTGGAACGGCGGTACTGCTTTTAC 300
 QY 301 CCCAGCCCCGAGATGCGGAGGCTTACAGGCTTCACTGCTCATGTTATTTAGTGT 360
 Db 301 CCCAGCCCCGAGATGCGGAGGCTTACAGGCTTCACTGCTCATGTTATTTAGTGT 360
 QY 361 TGGCTTAGCAGAAAGCTTCTGCTCATCTTTCGCGGAGATCCGTGGCACTCGCGTGG 420
 Db 361 TGGCTTAGCAGAAAGCTTCTGCTCATCTTTCGCGGAGATCCGTGGCACTCGCGTGG 420
 QY 421 TTTGTTGGTGAAGTTCTTCTGAGTCTGTCTATAGGCGGAGAAATTTGAGCTGAC 480
 Db 421 TTTGTTGGTGAAGTTCTTCTGAGTCTGTCTATAGGCGGAGAAATTTGAGCTGAC 480
 QY 481 TCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACCAACAATCTTCAAAAGCTTACG 540
 Db 481 TCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACCAACAATCTTCAAAAGCTTACG 540
 QY 541 CAGCGCGGCTTACAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 541 CAGCGCGGCTTACAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 601 TCACAGAGAACCCAGTGCATCAGCTGAAGAGCAATTTGACTTCAACAGACATTTAC 660
 Db 601 TCACAGAGAACCCAGTGCATCAGCTGAAGAGCAATTTGACTTCAACAGACATTTAC 660
 QY 661 GGGCTGTAAAGAAATTACGCCGAGAGTACGCAACGCACTGGAAGAGGGCTCCG 720
 Db 661 GGGCTGTAAAGAAATTACGCCGAGAGTACGCAACGCACTGGAAGAGGGCTCCG 720
 QY 721 ACCCAAGTCTTACCTGAGGAGAAAGTTTCAACCGAGTACGCTTGGGCTTACAC 780
 Db 721 ACCCAAGTCTTACCTGAGGAGAAAGTTTCAACCGAGTACGCTTGGGCTTACAC 780

QY 781 AGTACCACTGGAGGAGACATAAGCTTGGGCAAGCTATGGAGGCTTCTGCTTGGC 840
 Db 781 AGTACCACTGGAGGAGACATAAGCTTGGGCAAGCTATGGAGGCTTCTGCTTGGC 840
 QY 841 TCTCTTCCAAAGTGTCTTCCAGCGCCGCGCTTACGAGAGCTTGGCACTGCTGA 900
 Db 841 TCTCTTCCAAAGTGTGTCTTCCAGCGCCGCGCTTACGAGAGCTTGGCACTGCTGA 900
 QY 901 CCACGGAACCTTGGGCTTCTTGGGAGTGTGGCTTGGCTTCCANTCTTAGAGTGGCC 960
 Db 901 CCACGGAACCTTGGGCTTCTTGGGAGTGTGGCTTGGCTTCCANTCTTAGAGTGGCC 960
 QY 961 TCTGCGCGCTTCCGCTTCCAGGCTCTCCGCGCTACCACTAGTACAGGCGGCTTGG 1020
 Db 961 TCTGCGCGCTTCCGCTTCCAGGCTCTCCGCGCTACCACTAGTACAGGCGGCTTGG 1020
 QY 1021 TCACGCTGACAAACCGCGCTCTGTGCTTCTTCTTGGAGGAGGCGCTGTGAGTCT 1080
 Db 1021 TCACGCTGACAAACCGCGCTCTGTGCTTCTTCTTGGAGGAGGCGCTGTGAGTCT 1080
 QY 1081 ATGTTGCGCCAGCGCTTCTGCACTTCTGGAACCAAGCGCAAGAGCTGAGCAG 1140
 Db 1081 ATGTTGCGCCAGCGCTTCTGCACTTCTGGAACCAAGCGCAAGAGCTGAGCAG 1140
 QY 1141 AGAGAGGAGGCTCACCTTTATCTTGGGAGCCCACTGACAAAGAGGCGCTTCCAG 1200
 Db 1141 AGAGAGGAGGCTCACCTTTATCTTGGGAGCCCACTGACAAAGAGGCGCTTCCAG 1200
 QY 1201 ACTTAAATGATACCACTTATCTGTGAGGAGGAGCCCAATCTGACTCTTCCGCG 1260
 Db 1201 ACTTAAATGATACCACTTATCTGTGAGGAGGAGCCCAATCTGACTCTTCCGCG 1260
 QY 1261 TGGGACATGAGGCGCGGAGACAGTCCGCAAGGCTGCGGAGAGAGCTCCAGAA 1320
 Db 1261 TGGGACATGAGGCGCGGAGACAGTCCGCAAGGCTGCGGAGAGAGCTCCAGAA 1320
 QY 1321 GGGCACTGAGCGCTGTCGCGGAGAGCTTGGACATCCGCAAGCAACAGGAAAGTCT 1380
 Db 1321 GGGCACTGAGCGCTGTCGCGGAGAGCTTGGACATCCGCAAGCAACAGGAAAGTCT 1380
 QY 1381 TGGGCGGATCTGTAATAAACCCTTTTCTTTTGTGTTTTT 1421
 Db 1381 TGGGCGGATCTGTAATAAACCCTTTTCTTTTGTGTTTTT 1421

RESULT 2
 ADL06425
 ID ADL06425 standard; cDNA; 1450 BP.
 XX
 AC ADL06425;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #5.
 XX
 KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
 XX cancer; cytostatic; gene; ss.
 OS Homo sapiens.
 XX
 PN WO2004016225-A2.
 XX
 PD 26-FEB-2004.
 XX
 PF 19-AUG-2003; 2003WO-US025892.
 XX
 PR 19-AUG-2002; 2002US-0404809P.
 PR 21-AUG-2002; 2002US-0405645P.
 PR 23-SEP-2002; 2002US-0413192P.
 PR 15-OCT-2002; 2002US-0419008P.
 PR 15-NOV-2002; 2002US-0426847P.
 PR 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.
 XX Desauvage PJ, Frantz G, Hillan KJ, Polakie P, Polson A, Smith V,
 XX Spencer SD, Wu TD, Zhang Z;
 XX WPI: 2004-257144/24.
 XX P-PSDB; ADL06505.
 XX
 XX New antibody that binds to a tumor-associated antigenic target (TAT)
 XX polypeptide, useful for preparing a composition for diagnosing or
 XX treating cancer.
 XX
 XX Claim 1; SEQ ID NO 5; 319pp; English.
 XX
 XX The present invention relates to the isolation of human tumour-associated
 XX antigenic target (TAT) polynucleotide and TAT polypeptide sequences. Also
 XX disclosed is an antibody that binds to a TAT polypeptide. The antibody is
 XX a monoclonal antibody, an antibody fragment, a chimeric antibody or a
 XX humanised antibody. It is conjugated to a growth inhibitory agent. It is
 XX produced in bacteria or in CHO cells and induces death of a cell to which
 XX it binds. The antibody is useful for preparing a composition for
 XX diagnosing or treating tumours and cancer. The present sequence
 XX represents a human TAT cDNA sequence of the invention.
 XX
 XX Sequence 1450 BP; 279 A; 475 C; 398 G; 297 T; 0 U; 1 Other:
 XX
 XX Query Match 100.0%; Score 1421; DB 12; Length 1450;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 1421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 GAACCGAAGAAAGTACGCGCTACAGACAGTGAAGAAATGTTCCGTCGCGCGGTAGAAAA 60
 XX 9 GAACCGAAGAAAGTACGCGCTACAGACAGTGAAGAAATGTTCCGTCGCGCGGTAGAAAA 68
 XX 61 CTCTGTCGATGACCAACCCGAGAGGTGAGAGAGCCCACTCCAGCTTCTTAAAGGA 120
 XX 69 CTCTGTCGATGACCAACCCGAGAGGTGAGAGAGCCCACTCCAGCTTCTTAAAGGA 128
 XX 121 GAGGTGAGAGACTCAGACTTCAACAGCCCACTCGGTCCAGCCTTGTAGCAAGAGAG 180
 XX 129 GAGGTGAGAGACTCAGACTTCAACAGCCCACTCGGTCCAGCCTTGTAGCAAGAGAG 188
 XX 181 CCAAGAGAGCGCTCTCCGCGGTCCAGAGAGCCCACTGCTTGTGCTTCCGCGCTG 240
 XX 189 CCAAGAGAGCGCTCTCCGCGGTCCAGAGAGCCCACTGCTTGTGCTTCCGCGCTG 248
 XX 241 CGTGAGAGACTCGCGCGGTGAGAGAGCCCACTGCTTGTGAGAGCGGCTTGTGAG 300
 XX 249 CGTGAGAGACTCGCGCGGTGAGAGAGCCCACTGCTTGTGAGAGCGGCTTGTGAG 308
 XX 301 CCGAGAGCGCGAGTGCAGAGAGCTTCAAGCTTCACTGCTATGTTAGTGT 360
 XX 309 CCGAGAGCGCGAGTGCAGAGAGCTTCAAGCTTCACTGCTATGTTAGTGT 368
 XX 361 TGGCTTAGAGAGAGCTTCTCTGCTCATCTTCCGCGGAGTCCGTGCGACTCGGCTG 420
 XX 369 TGGCTTAGAGAGAGCTTCTCTGCTCATCTTCCGCGGAGTCCGTGCGACTCGGCTG 428
 XX 421 TTTGGTTGGAGAGGTTCTTCTAGTCTGTTCAATAGGCGAGAAATGTGGCTGTGACT 480
 XX 429 TTTGGTTGGAGAGGTTCTTCTAGTCTGTTCAATAGGCGAGAAATGTGGCTGTGACT 488
 XX 481 TCAGTGAGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
 XX 489 TCAGTGAGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 548
 XX 541 CAGGCGCGTTACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 XX 549 CAGGCGCGTTACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 608
 XX 601 TCACAGGAGACCCGATGATCAGTGAAGAGACATTAAGTGAAGAGAGGTTCACT 660

Db 609 TCACAGGAGACCCGATGATCAGTGAAGAGACATTAAGTGAAGAGAGGTTCACT 668
 Qy 661 GCGCTCTGAAGAAGATTACGCGCGAGTACGCGAGACACTGAGAGAGGCTGCGCG 720
 Db 669 GCGCTCTGAAGAAGATTACGCGCGAGTACGCGAGACACTGAGAGAGGCTGCGCG 728
 Qy 721 ACCAGTGTCTTACCTGCGGAGAAAGTTCAACCGAGTACGCTTCCGCTGTACACC 780
 Db 729 ACCAGTGTCTTACCTGCGGAGAAAGTTCAACCGAGTACGCTTCCGCTGTACACC 788
 Qy 781 AGTACCACTGCGGAGACACTACGCTCGGCAAGCTAAGGTGCTTCTGCTTCTGCG 840
 Db 789 AGTACCACTGCGGAGACACTACGCTCGGCAAGCTAAGGTGCTTCTGCTTCTGCG 848
 Qy 841 TCTCTTCAACGTTGCTCTTCAAGCGCGCGCGCTTCAAGAGGCTGAGCTGTA 900
 Db 849 TCTCTTCAACGTTGCTCTTCAAGCGCGCGCTTCAAGAGGCTGAGCTGTA 908
 Qy 901 CCAACGAGACTTTCGCGCTTTCGAGGCTTTCGCTTTCGCTTTCATCTTACGCTGCG 960
 Db 909 CCAACGAGACTTTCGCGCTTTCGAGGCTTTCGCTTTCGCTTTCATCTTACGCTGCG 968
 Qy 961 TCTGCGCTTCGCGCTTTCGAGGCTTTCGCTTTCGCTTTCATCTTACGCTGCG 1020
 Db 969 TCTGCGCGCTTCGCGCTTTCGAGGCTTTCGCTTTCGCTTTCATCTTACGCTGCG 1028
 Qy 1021 TCAAGCTGAGACCCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCATCTTACG 1080
 Db 1029 TCAAGCTGAGACCCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCATCTTACG 1088
 Qy 1081 ATGTTGCGCGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAG 1140
 Db 1089 ATGTTGCGCGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAG 1148
 Qy 1141 AGAGAGGCGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAG 1200
 Db 1149 AGAGAGGCGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAGGCTTTCGAG 1208
 Qy 1201 ACTTAAATGTTATACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1260
 Db 1209 ACTTAAATGTTATACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1268
 Qy 1261 TGGAGATGCGAGAGCGGAGAGAGTGCAGGAGCTTGGGCGAGAGAGTCTCAGAA 1320
 Db 1269 TGGAGATGCGAGAGCGGAGAGAGTGCAGGAGCTTGGGCGAGAGAGTCTCAGAA 1328
 Qy 1321 GGGCACTGAGCGCTGCTGCGGAGAGCTTGGGCGAGAGAGTCTCAGAA 1380
 Db 1329 GGGCACTGAGCGCTGCTGCGGAGAGCTTGGGCGAGAGAGTCTCAGAA 1388
 Qy 1381 TGGGCGAGTGTAAATAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1421
 Db 1389 TGGGCGAGTGTAAATAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1429
 XX
 XX RESULT 3
 XX ABL90535
 XX ID ABL90535 standard; cDNA; 1474 BP.
 XX
 XX AC ABL90535;
 XX XX
 XX DT 24-MAY-2002 (first entry)
 XX XX
 XX DE Human polynucleotide SEQ ID NO 1097.
 XX XX
 XX XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX XX antiallergic; hepatotropic; antidiabetic; antifungal; antiparasitic;
 XX XX antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic;
 XX XX caducant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX XX neurological disease; infection; human; secreted protein; gene; ss.
 XX OS Homo sapiens.
 XX XX

XX AA94623;
 AC 11-JAN-2001 (first entry)
 XX Human CASB618 coding sequence.
 XX Epitope: human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
 KM colon; autoimmune disease; HLA_A0201; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 259..1221
 FT /tag= a
 FT /product= "Human CASB618"
 PN W020053748-A2.
 PD 14-SEP-2000.
 PF 09-MAR-2000; 2000MO-EP002048.
 PR 11-MAR-1999; 99GB-00005607.
 PR 01-SEP-1999; 99GB-00020590.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Bruck CEW, Casasart J, Coche T, Vinals Y De Baesolac;
 DR WPI: 2000-572268/53.
 DR P-PSDB; AAB26325.
 XX New human CASB618 polypeptide, useful as a vaccine for prophylactic and
 PT therapeutic treatment of cancers, particularly ovarian or colon cancer,
 PT autoimmune diseases and related conditions.
 XX Claim 13; Page 61; 76pp; English.
 CC The present sequence is the coding sequence of human CASB618 protein. The
 CC gene for human CASB618 is thought to be located on chromosome 15. The
 CC protein encoded by the present sequence and epitopes of the CASB618
 CC protein (see AAB26327 to AAB6339) are useful in diagnosing the occurrence
 CC of tumour cells and in vaccines for prophylactic and therapeutic
 CC treatment of cancers, particularly ovarian or colon cancer, autoimmune
 CC diseases and related conditions
 CC
 SQ Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other;
 Query Match 99.4%; Score 1413; DB 3; Length 1441;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 CCGCATGCCGACAGGCTTACAGCGTTCCATGCTCATGTTATTTAGTGTGTTGGCTCTA 368
 DB 301 CGGCATGCCGACAGGCTTACAGCGTTCCATGCTCATGTTATTTAGTGTGTTGGCTCTA 360
 QY 369 GCAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGCCCATCTGGCGTGGTTGGTTG 428
 DB 361 GCAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGCCCATCTGGCGTGGTTGGTTG 420
 QY 429 GTGAGAGTCTTCTGCTCATCTTGGCGGAGATCCGTGCCCATCTGGCGTGGTTGGTTG 488
 DB 421 GTGAGAGTCTTCTGCTCATCTTGGCGGAGATCCGTGCCCATCTGGCGTGGTTGGTTG 480
 QY 489 GAATGCTTGTGGGTACAGTGAACCAACCACTCTCAAAAGCTTACAGCGAGCGCGC 548
 DB 481 GAATGCTTGTGGGTACAGTGAACCAACCACTCTCAAAAGCTTACAGCGAGCGCGC 540
 QY 549 GTTACAGCCCGTGTGGTGTCTGTGGGCTGTGGAGGCAATTATTTACCTCACAGGG 608
 DB 541 GTTACAGCCCGTGTGGTGTCTGTGGGCTGTGGAGGCAATTATTTACCTCACAGGG 600
 QY 609 ACCCGATGTCATGAGTGAACGAGACCACTTGAACAAGAGAGTTCACCTGGCGTCTG 668
 DB 601 ACCCGATGTCATGAGTGAACGAGACCACTTGAACAAGAGAGTTCACCTGGCGTCTG 660
 QY 669 AAAGAAATTTACGCGCGAGTACGCGAAACGACTGAGAGAGGAGGCTGCGGACCCAGTG 728
 DB 661 AAAGAAATTTACGCGCGAGTACGCGAAACGACTGAGAGAGGAGGCTGCGGACCCAGTG 720
 QY 729 CTCTACCTGGCGGAGAAATTTACACGAGTACGCTTGGCGCTGTACCAACAGTACAC 788
 DB 721 CTCTACCTGGCGGAGAAATTTACACGAGTACGCTTGGCGCTGTACCAACAGTACAC 780
 QY 789 CTGGCGGAGACACTACGCTTGGCGGAGTATGAGTATGAGTATGAGTATGAGTATGAG 848
 DB 781 CTGGCGGAGACACTACGCTTGGCGGAGTATGAGTATGAGTATGAGTATGAGTATGAG 840
 QY 849 AAGCTCTGCTCTCAACGCGCGGCTCTCAAGAGGCTGTGACCTGTGACCAACGGA 908
 DB 841 AAGCTCTGCTCTCAACGCGCGGCTCTCAAGAGGCTGTGACCTGTGACCAACGGA 900
 QY 909 GCTTGGCGCTCTTGGCGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 968
 DB 901 GCTTGGCGCTCTTGGCGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 960
 QY 969 CTCCGCTTGAAGCTCTTCCGCGCTCAACCTCAATAGGCGCGCTTGGGTCACGCTG 1028
 DB 961 CTCCGCTTGAAGCTCTTCCGCGCTCAACCTCAATAGGCGCGCTTGGGTCACGCTG 1020
 QY 1029 GCAACCGGCGTCTGTGCTCTTCTCTGAGAGGCGGTGTGAGTCTCAAGTATGTTGCG 1088
 DB 1021 GCAACCGGCGTCTGTGCTCTTCTCTGAGAGGCGGTGTGAGTCTCAAGTATGTTGCG 1080
 QY 1089 CCCAGGCTCTTGTGCAACCTTCTGTGAACAAAGCGCCAAAGACTGTGACCCAGAGAGAGG 1148
 DB 1081 CCCAGGCTCTTGTGCAACCTTCTGTGAACAAAGCGCCAAAGACTGTGACCCAGAGAGAGG 1140
 QY 1149 GGTCTACCTTATTTCTCTGGCGACCACTGTGACAAAGAGGCGCTCTCCAGACTTAAA 1208
 DB 1141 GGTCTACCTTATTTCTCTGGCGACCACTGTGACAAAGAGGCGCTCTCCAGACTTAAA 1200
 QY 1209 TGTATCAACCACTAACCTGTGAGGAGGAGCCCAATCTGGAATCTTCCCGGCTTGGAGAT 1268
 DB 1201 TGTATCAACCACTAACCTGTGAGGAGGAGCCCAATCTGGAATCTTCCCGGCTTGGAGAT 1260
 QY 1269 CGCAGGCGGGAAGAGTGTGCGCGACAGGCTGTGGGCGAGAGACTCCAGAAAGGCACTG 1328
 DB 1261 CGCAGGCGGGAAGAGTGTGCGCGCGACAGGCTGTGGGCGAGAGACTCCAGAAAGGCACTG 1320
 QY 1329 AGGCTGTGTGGCGGAGGCTGTGAGATCCGAGGAGCAACGAGGAAAGTCTCTGGGCGCA 1388
 DB 1321 AGGCTGTGTGGCGGAGGCTGTGAGATCCGAGGAGCAACGAGGAAAGTCTCTGGGCGCA 1380
 QY 1389 TCTGTAATTAACCTTTTCTTTTGTGTTT 1421

DB 1381 TCTGTAAATTAACCTTTTCTTTCTTTT 1413

RESULT 5
ADD192228
ID ADD192228 standard; cDNA; 1491 BP.
XX ADD192228;
AC
XX 15-JAN-2004 (first entry)
XX
XX
DE Human cDNA from secreted protein gene 45.
XX
XX human secreted protein; cytostatic; antibacterial; virucide;
KM neoprotective; gynaecological; gastrointestinal-gen; cardiant;
KM cardiovascular-gen; nephrotoxic; antiinflammatory; muscular-gen;
KM respiratory-gen; immunosuppressive; cerebroprotective; vasotropic;
KM neurotropic; antiallergic; cancer; bacterial infection; viral infection;
KM neutral disorder; immune system disorder; blood disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder;
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM inflammatory disorder; proliferative disorder; human; ss; gene.
XX
OS Homo sapiens.
XX MO2003052377-A2.
XX
XX 26-JUN-2003.
XX
XX 06-NOV-2002; 2002WO-US035606.
XX PF
XX 07-NOV-2001; 2001US-0331046P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX MPI; 2003-533050/50.
XX P-PSDB; ADD19303.
XX
XX
XX New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX
XX
XX Claim 1; SEQ ID NO 55; 554pp; English.
XX
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject,
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities: cytostatic, antibacterial,
CC virucide, neuroprotective, gynaecological, gastrointestinal-gen,
CC Cardiant, Cardiovascular-gen, Nephrotoxic, Antiinflammatory, Muscular-
CC Gen, Respiratory-gen, Immunosuppressive, Cerebroprotective, Vasotropic,
CC Neurotropic, Antiallergic. The methods and compositions of the present
CC invention are useful for diagnosing, treating, preventing and/or

CC prognosticating disorders related to the novel polypeptides, such as
CC cancer, bacterial or viral infections, and neural, immune system, blood,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, inflammatory or proliferative disorders (many examples of these
CC diseases and disorders are given in the specification). The present
CC sequence encodes a novel secreted protein of the invention.

XX Sequence 1491 BP; 293 A; 485 C; 410 G; 302 T; 0 U; 1 Other;

Query Match 99.4%; Score 1412; DB 10; Length 1491;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GAAAGTAACGGCTACAGACAGTGAAGAAATGTTGCTGCGCGGTAGAAAACTCTGTC 67
DB 43 GAAAGTAACGGCTACAGACAGTGAAGAAATGTTGCTGCGCGGTAGAAAACTCTGTC 102
QY 68 GGTACCAACCCGAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAA CGAGAGTGTC 127
DB 103 GGTACCAACCCGAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAA CGAGAGTGTC 162
QY 128 AGGACTCAGACTTTCACACAGCCCACTGCGTCCAGCTTGTACGCAAGAGAGCCAAAGA 187
DB 163 AGGACTCAGACTTTCACACAGCCCACTGCGTCCAGCTTGTACGCAAGAGAGCCCAAGA 222
QY 188 CGCGCTCTCCGCGGTCCAGAGAGCCCACTTGTGCTGCTGCGCGCTGCGTGCAG 247
DB 223 CGCGCTCTCCGCGGTCCAGAGAGCCCACTTGTGCTGCTGCGCGCTGCGTGCAG 282
QY 248 CACTGCGCGCGGTGAGAGCATGACCTGTGGAACGGCGTACTGCTTTTAA CCGCAGGC 307
DB 283 CACTGCGCGCGGTGAGAGCATGACCTGTGGAACGGCGTACTGCTTTTAA CCGCAGGC 342
QY 308 CCGGCAATGCGGAGAGCTTCAAGCTTCACTGCTCATCCGTTATTCAGATGTTTTGGCTCT 367
DB 343 CCGGCAATGCGGAGAGCTTCAAGCTTCACTGCTCATCCGTTATTCAGATGTTTTGGCTCT 402
QY 368 AGAGCAAGCTTCTGCTCATCTGCTGCGGAGATCCGAGGCACTCGCGCTGTTTTGGTT 427
DB 403 AGAGCAAGCTTCTGCTCATCTGCTGCGGAGATCCGAGGCACTCGCGCTGTTTTGGTT 462
QY 428 GGTGAGAGTTCTTCTCAGTCTGTTCTATAGCGCAGAAATTTGGCTGTGCATTTCAAGTGC 487
DB 463 GGTGAGAGTTCTTCTCAGTCTGTTCTATAGCGCAGAAATTTGGCTGTGCATTTCAAGTGC 522
QY 488 AGAATGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 547
DB 523 AGAATGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 582
QY 548 CGTTACAGCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
DB 583 CGTTACAGCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
QY 608 GACCCAGTGCATCAGCTGAACGAGACATTTGATTCACAGACAGTTCACCTGCGTCT 667
DB 643 GACCCAGTGCATCAGCTGAACGAGACATTTGATTCACAGACAGTTCACCTGCGTCT 702
QY 668 GAAAGGAATTAACGCGCGGAGTACCGCAACGACCTGAGGAGGGCTCCCGAACCTCAGT 727
DB 703 GAAAGGAATTAACGCGCGGAGTACCGCAACGACCTGAGGAGGGCTCCCGAACCTCAGT 762
QY 728 GCTCTACCTGCGGAGGAATTTCAACACGAGTGAACCTTGCAGGCTTATACCAACAGTACCA 787
DB 763 GCTCTACCTGCGGAGGAATTTCAACACGAGTGAACCTTGCAGGCTTATACCAACAGTACCA 822
QY 788 CTTGCGGAGCACTACGCTTGCGCAAGCTATAGGTGAGGCTTCTGCTGCTCTCTC 847
DB 823 CTTGCGGAGCACTACGCTTGCGCAAGCTATAGGTGAGGCTTCTGCTGCTCTCTC 882
QY 848 CAACTGCTGCTCTCAGCGCGCGCGCTTACGGAAGGCTGCGCATCTGACCAACCG 907
DB 883 CAACTGCTGCTCTCAGCGCGCGCGCTTACGGAAGGCTGCGCATCTGACCAACCG 942


```

Db 554 CAGCGCGGCTTACAGCCCGGTGCTGCTGCTGCGGCGTGGAGGCGATTATATTAAC 613
QY 601 TCACAGGAGCCCGAGTGCATCACTGAACAGAACATTTGACTAACAGACATTTCACT 660
Db 614 TCACAGGAGCCCGAGTGCATCACTGAACAGAACATTTGACTAACAGACATTTCACT 673
QY 661 GCGCTGTAAGAGATTAAGCGCGGAGTACCGAGAACGCACTGAGAAAGGCGTCCG 720
Db 674 GCGCTGTAAGAGATTAAGCGCGGAGTACCGAGAACGCACTGAGAAAGGCGTCCG 733
QY 721 ACCCACTGCTCTACCTGCGGAGAGTTCAACCGAGTACGCTTGGCGCTGTACACC 780
Db 734 ACCCACTGCTCTACCTGCGGAGAGTTCAACCGAGTACGCTTGGCGCTGTACACC 793
QY 781 AGTACCACTGCGGAGCACTACGCTCGGCGAGCTAGTGGGAGGCTTGTCTTGGC 840
Db 794 AGTACCACTGCGGAGCACTACGCTCGGCGAGCTAGTGGGAGGCTTGTCTTGGC 853
QY 841 TCCTCTCAACGCTGCTCTTCCAGCCGCGCTCTTACCGAGGCTTGGCACTGCTGA 900
Db 854 TCCTCTCAACGCTGCTCTTCCAGCCGCGCTCTTACCGAGGCTTGGCACTGCTGA 913
QY 901 CCACCGGAGCTTGGGCTTTTGGGAGTCTTGGCTTGGCTTCACTTCACTGCTGCGC 960
Db 914 CCACCGGAGCTTGGGCTTTTGGGAGTCTTGGCTTGGCTTCACTTCACTGCTGCGC 973
QY 961 TCAGCCGCTCGGCTAGGCTCTTCCGCGCTACCACTGATAGGAGCGGCTTGGAG 1020
Db 974 TCAGCCGCTCGGCTAGGCTCTTCCGCGCTACCACTGATAGGAGCGGCTTGGAG 1033
QY 1021 TCAGCTGCGAACCGGCGCTTCTGCTCTTCTTCCGAGGCGGCTGATGCTTCCAGT 1080
Db 1034 TCAGCTGCGAACCGGCGCTTCTGCTCTTCTTCCGAGGCGGCTGATGCTTCCAGT 1093
QY 1081 ATGTTGGCGCCAGGCTTTTGGACCTTCTTGGACCAAAAGCGCCAGAGCTGC-AGCCAG 1139
Db 1094 ATGTTGGCGCCAGGCTTTTGGACCTTCTTGGACCAAAAGCGCCAGAGCTGCAG 1153
QY 1140 GAGAGAGGCGGCTCACTTATCTCTGCGGAGCCCACTGACCAAGAGGCGGCTTCCCA 1199
Db 1154 GAGAGAGGCGGCTCACTTATCTCTGCGGAGCCCACTGACCAAGAGGCGGCTTCCCA 1212
QY 1200 GAC-TTAAATATGATCACACTTACCTGTGAGGAGGAGCCCAATCTGACTCTTCCCGC 1258
Db 1213 GACTTAAATATGATCACACTTACCTGTGAGGAGGAGCCCAATCTGACTCTTCCCGC 1272
QY 1259 CTGAGGACATTCGAGGCGCGGAGAGGAGTCCGCGGAGGCTGCGGAGAGAGCTTCAAG 1318
Db 1273 CTGAGGACATTCGAGGCGCGGAGAGGAGTCCGCGGAGGCTGCGGAGAGAGCTTCAAG 1332
QY 1319 AAGGAGCACTGAGGCTGCTGCGCGAGAGGCTCGGACATCCGAGGAGCAAGGAAAGTCT 1378
Db 1333 AAGGAGCACTGAGGCTGCTGCGCGAGAGGCTCGGACATCCGAGGAGCAAGGAAAGTCT 1392
QY 1379 CCGTGGGCGATCTGTAAATTAACCTTTTCTTTTGTGTTTTT 1421
Db 1393 CCGTGGGCGATCTGTAAATTAACCTTTTCTTTTGTGTTTTT 1435

```

RESULT 7
 AB273640
 ID AB273640 standard; cDNA, 1460 BP.
 AC AB273640;
 XX
 XX 12-MAY-2003 (first entry)
 DE Secreted protein-encoding gene 360 cDNA clone HUFCLJ1, SEQ ID NO:370.
 XX
 XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;

KW drug screening; chromosome identification; chromosome mapping;
 KW cytotoxic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antineoplastic; vulnery; chromosome 9p21; gene; ss.
 OS Homo sapiens.
 PN MO200277013-A2.
 PD 03-OCT-2002.
 PF 26-MAR-2002; 2002WO-US009370.
 PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 PI WPI; 2003-040578/03.
 DR P-PSDB; ABR01306.
 XX
 XX New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX
 PS Claim 21; Page 1345; 2474pp; English.
 XX
 XX AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC AB273698-AB274687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prothrombin activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein-encoding cDNA clone of the invention
 XX
 SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match 96.8%; Score 1375.8; DB 8; Length 1460;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
 QY 1 GAACAGAGAAATGATGAGTGAAGAAATAGTTGCTGCGCGGTAAAGAAA 60
 Db 14 GAACAGAGAAATGATGAGTGAAGAAATAGTTGCTGCGCGGTAAAGAAA 73
 QY 61 CTGTGTGGTACCAACCCGAGGCTTGAAGAGCCCACTTCAAGCTTCTTAAAGGA 120
 Db 74 CTGTGTGGTACCAACCCGAGGCTTGAAGAGCCCACTTCAAGCTTCTTAAAGGA 133
 QY 121 GAGGTGCAAGACTGAGACTTCCAGCCCACTGCGGCTTGAAGAGAGAG 180
 Db 134 GAGGTGCAAGACTGAGACTTCCAGCCCACTGCGGCTTGAAGAGAGAG 193
 QY 181 CCAAGAGAGCGCTTCCGCGCTCAGAGAGCCCAAGCTTGTGCTGCTCCGCGCTG 240
 Db 194 TCAAGAGAGCGCTTCCGCGCTCAGAGAGCCCAAGCTTGTGCTGCTCCGCGCTG 253

QY 241 CGTGCAGCATCTGGCGCGGTGAGCATGACCTGTGTGAACGGCGCTGACCTTTTATAC 300
 Db 254 CGTGCAGCATCTGGCGCGGTGAGCATGACCTGTGTGAACGGCGCTGACCTTTTATAC 313
 QY 301 CCCAGCCCCGCGAGCGCGGCTTCAAGGCTTCACTGCTCATGTTATTTAGTGT 360
 Db 314 CCCAGCCCCGCGAGCGCGGCTTCAAGGCTTCACTGCTCATGTTATTTAGTGT 373
 QY 361 TGGCTTACAGAACTTCTGCTCATCTTGGCGGGGATCCGTGGCGCATCGCGCTGGT 420
 Db 374 TGGCTTACAGAACTTCTGCTCATCTTGGCGGGGATCCGTGGCGCATCGCGCTGGT 433
 QY 421 TTTGGTTGAGAGTTCTTCTCACTCTGTTCATAGCGCGAATTTGGCTGTGCACT 480
 Db 434 TTTGGTTGAGAGTTCTTCTCACTCTGTTCATAGCGCGAATTTGGCTGTGCACT 493
 QY 481 TCAGTCAGAAATGTTCTGTGGGTACAGTGAACCAACATCTTCAAAAGCTTCAGCG 540
 Db 494 TCAGTCAGAAATGTTCTGTGGGTACAGTGAACCAACATCTTCAAAAGCTTCAGCG 553
 QY 541 CAGCGCGCTTACAGCGCGCTGTGCTGTGGCGCTTGGCGCGGATTAATTAAC 600
 Db 554 CAGCGCGCTTACAGCGCGCTGTGCTGTGGCGCTTGGCGCGGATTAATTAAC 613
 QY 601 TCACAGGGAACCCAGTGCATCAGTGAACAGACATTTGACTAACAAGCAGTCACT 660
 Db 614 TCACAGGGAACCCAGTGCATCAGTGAACAGACATTTGACTAACAAGCAGTCACT 673
 QY 661 GGGCTGTGAAGAAATTACCGCGGAGTACGGAACGCACTGAGAAAGGGCTGCGG 720
 Db 674 GGGCTGTGAAGAAATTACCGCGGAGTACGGAACGCACTGAGAAAGGGCTGCGG 733
 QY 721 ACCGAGTGTCTTACCTGGCGGAGAAATTCACACCGAGTACCTTGGCGCTGTACAC 780
 Db 734 ACCGAGTGTCTTACCTGGCGGAGAAATTCACACCGAGTACCTTGGCGCTGTACAC 793
 QY 781 AGTACCACTGGGGGAGCACTACGCTCGGCGACGCTATGGGTGGCGCTTGTGCTGAG 840
 Db 794 AGTACCACTGGGGGAGCACTACGCTCGGCGACGCTATGGGTGGCGCTTGTGCTGAG 853
 QY 841 TCCCTTCACAGTGTGCTCTTCCAGCGCGCGCTTCAAGAGGCTTGGACCTGCTGA 900
 Db 854 TCCCTTCACAGTGTGCTCTTCCAGCGCGCGCTTCAAGAGGCTTGGACCTGCTGA 913
 QY 901 CCAACCGAGCTTGGCGCTCTTGGGGGTCTTGGCTTGGCTCATCTTACAGGTGCGC 960
 Db 914 CCAACCGAGCTTGGCGCTCTTGGGGGTCTTGGCTTGGCTCATCTTACAGGTGCGC 973
 QY 961 TCTGCGCGCTCGGCTTACGCTCCCGCGCTCACTCACTGATCGCGCGCTTGTGG 1020
 Db 974 TCTGCGCGCTCGGCTTACGCTCCCGCGCTCACTCACTGATCGCGCGCTTGTGG 1033
 QY 1021 TCAAGCTGAGCAACGCGCTCTGTGCTCTTCCCTCGAGGGGCGCTGTGATCTCCAGT 1080
 Db 1034 TCAAGCTGAGCAACGCGCTCTGTGCTCTTCCCTCGAGGGGCGCTGTGATCTCCAGT 1093
 QY 1081 ATGTTCGCGCAAGGCTCTTGGCACTTGTGACCAAGCGCTCAAGAGATGTC-AGCAG 1139
 Db 1094 ATGTTCGCGCAAGGCTCTTGGCACTTGTGACCAAGCGCTCAAGAGATGTC-AGCAG 1153
 QY 1140 GAGAGAGGGGCTCACTTATCTGTGGGAGCCCACTGACCAAGAGGCGCTGCCA 1199
 Db 1154 GAGAGAGGGGCTCACTTATCT-GGGAGCCCACTGACCAAGAGGCGCTGCCA 1212
 QY 1200 GAC-TTAAATGTATCACTAATCTGTGAGGGGAGCCCAATCTGACTCTTCCCGC 1258
 Db 1213 GACTTTAAATGTATCACTAATCTGTGAGGGGAGCCCAATCTGACTCTTCCCGC 1272
 QY 1259 CTTGGGACATCGCAGCGCGGAGAGAGAGTCCCGCAGGCTTGGGCGCAGAGAGTCCAGG 1318
 Db 1273 CTTGGGACATCGCAGCGCGGAGAGAGAGTCCCGCAGGCTTGGGCGCAGAGAGTCCAGG 1332
 QY 1319 AAGGGCACTAGAGCGCTGTGGCGGAGAGGCTTCGAGCATTCGAGCAGGACAGGAAAGTCT 1378

Db 1333 AAGGGCACTAGAGCGCTGTGGCGCGAGGCTCGGACATCCGAGGACCAAGGAAAGTCT 1392
 QY 1379 CTTGGGCGAGCTGTAAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1421
 Db 1393 CTTGGGCGAGCTGTAAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1435
 RESULT 8
 ADCC20289
 ID ADCC20289 standard; DNA; 1460 BP.
 XX
 AC ADCC20289;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human secreted protein coding sequence #228.
 XX
 KW gene therapy; human; secreted protein; haemopoietic disorder;
 KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
 KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
 KW leukemia; wound healing; epithelial cell proliferation disorder;
 KW immune disorder; autoimmune disorder; asthmatic disorder;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
 KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200292787-A2.
 XX
 PD 21-NOV-2002.
 XX
 PF 26-MAR-2002; 2002MO-US009257.
 XX
 PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-129287/12.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating hematopoietic or hematologic disorders, e.g.
 PT anemia or hemophilia.
 XX
 PS Claim 1, SEQ ID NO 238; 1512pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC secreted proteins. The DNA and protein sequences of the invention are
 CC useful for detecting, preventing, diagnosing, prognosticating, treating
 CC or ameliorating; hematopoietic or haematological disorders (e.g. anaemia
 CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
 CC and Crohn's disease); neoplastic disease (e.g. cancer and leukemia);
 CC wound healing and disorders of epithelial cell proliferation; immune
 CC disorders (e.g. autoimmune disorders and asthmatic disorders);
 CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
 CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
 CC and gastrointestinal disorders (e.g. duodenal ulcers and
 CC gastroenteritis). The present DNA sequence encodes a human secreted
 CC protein of the invention.
 XX
 SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match 96.84; Score 1375.8; DB 10; Length 1460;
 Best Local Similarity 99.38; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

QY 1 GAACAGGAAGTAACGGCTACAGACAGTGAAGAAATGTTGCTCGCGGCTAGAAAAA 60

Db 14 GAACGAGAAAGTAAACGCTACAGACAGTGAAGAAATATGTTGCTGCGCGGCTAGAAAA 73
Qy 61 CTCTGTGGTACCAACACCCAGAGGCTTTGAGAGAGCCCACTTCAGCTTCTCTTAACGA 120
Db 74 CTCTGTGGTACCAACCCAGAGGCTTTGAGAGAGCCCACTTCAGCTTCTCTTAACGA 133
Qy 121 GAGGTGACAGGACTCAGACTTCAACAGCCCACTCGTCCAGGCTTTGTAAGCAAGAGAG 180
Db 134 GAGGTGACAGGACTCAGACTTCAACAGCCCACTCGTCCAGGCTTTGTAAGCAAGAGAG 193
Qy 181 CCAAGAGAGCGGCTCTCCGCGTCAAGGCAAGCCCACTTGTGACTTGTGCTGCGCGCTG 240
Db 194 TCAAGGAGCGGCTCTCCGCGTCAAGGCAAGCCCACTTGTGACTTGTGCTGCGCGCTG 253
Qy 241 CGTGACAGCACTGGGCGGCGTGAAGATACCTGTGGAAGCGGCGTATGCTTTTACG 300
Db 254 CGTGACAGCACTGGGCGGCGTGAAGATACCTGTGGAAGCGGCGTATGCTTTTACG 313
Qy 301 CCAAGCGCGGCGATGCGGAGGCTTCAAGGCTTCACTGCTCATGTTATTTAGTGT 360
Db 314 CCAAGCGCGGCGATGCGGAGGCTTCAAGGCTTCACTGCTCATGTTATTTAGTGT 373
Qy 361 TGGCTTACAGAGACTTCTGCTCATCTTTCGCGGAGATCCGTTGCGCACTGCGCTG 420
Db 374 TGGCTTACAGAGACTTCTGCTCATCTTTCGCGGAGATCCGTTGCGCACTGCGCTG 433
Qy 421 TTTGTTGGTGAAGTCTTTCAGTCTGTTCATAGCCCAAAATTTGGTGGCTGCACT 480
Db 434 TTTGTTGGTGAAGTCTTTCAGTCTGTTCATAGCCCAAAATTTGGTGGCTGCACT 493
Qy 481 TCAGTGCAGATAGTTCGTGGGTACAGTGAACCAACATCTCAAGGCTTTCAGCG 540
Db 494 TCAGTGCAGATAGTTCGTGGGTACAGTGAACCAACATCTCAAGGCTTTCAGCG 553
Qy 541 CAGCGCGGCTTACAGCCCGTGTGCTGCTGCGGCTTGAAGGCGATTAATATTAC 600
Db 554 CAGCGCGGCTTACAGCCCGTGTGCTGCTGCGGCTTGAAGGCGATTAATATTAC 613
Qy 601 TCACAGGAGCCCGAGTGCATAGCTGAAGAGCACTTGAATCAACAGAGGTTACCT 660
Db 614 TCACAGGAGCCCGAGTGCATAGCTGAAGAGCACTTGAATCAACAGAGGTTACCT 673
Qy 661 GCGGTCTGAAGAAATTAACCGCGAGTACCGCAACGCACTGAGAAAGGCGCTCCG 720
Db 674 GCGGTCTGAAGAAATTAACCGCGAGTACCGCAACGCACTGAGAAAGGCGCTCCG 733
Qy 721 ACCAGTGTCTTACCTGGCGAGAGTTCAACCGAGTAGCCCTTGGCGCTGTACACC 780
Db 734 ACCAGTGTCTTACCTGGCGAGAGTTCAACCGAGTAGCCCTTGGCGCTGTACACC 793
Qy 781 AGTACCACTGGGAGGAGCACTAGGCTCGGCAAGGCTAGTAGGAGGCTTGTGAG 840
Db 794 AGTACCACTGGGAGGAGCACTAGGCTCGGCAAGGCTAGTAGGAGGCTTGTGAG 853
Qy 841 TCTCTTCAACGCTGCTCTCCAGCGCGCGCTTCAAGAGGCTTGTGAGCTGCTGA 900
Db 854 TCTCTTCAACGCTGCTCTCCAGCGCGCGCTTCAAGAGGCTTGTGAGCTGCTGA 913
Qy 901 CCAAGCGAGCTTGGCGCTTTCGGGAGTCTTGGCTTGGCTTCATCTCAAGCTGCGCG 960
Db 914 CCAAGCGAGCTTGGCGCTTTCGGGAGTCTTGGCTTGGCTTCATCTCAAGCTGCGCG 973
Qy 961 TCGGCGCGCTCGGCTAGGCTCTCGGCGCTACCACTAGTACGGGCGCGCTTTCG 1020
Db 974 TCGGCGCGCTCGGCTAGGCTCTCGGCGCTACCACTAGTACGGGCGCGCTTTCG 1033
Qy 1021 TCAAGCTGAGCAACCGGCTCTGCTGCTTCTCTGAGAGGCGCTGTGAGTCTCAAG 1080
Db 1034 TCAAGCTGAGCAACCGGCTCTGCTGCTTCTCTGAGAGGCGCTGTGAGTCTCAAG 1093
Qy 1081 ATGTTGGCGCGAGGCTTTCGCACTTTCGAGCAAAAGCGCAAGAGACTGC-AGCCAG 1139

Db 1094 ATGTTGCGCCAGCGCTCTTCGACCTTCTGAGACCAAGCCCAAGAGACTGCAAGCCAG 1153
Qy 1140 GAGAGAGGGGGCTCACTCTTATCTCTCGGAGACCACTGACAGAGGCGCTCCGA 1199
Db 1154 AAGAGGGGGGGCTCACTCTTATCT-GGAGACCACTGACAGAGAGGCGCTTCCCA 1212
Qy 1200 GAC-TTAAATGATACACCACTTACCTGTGAGGGGAGCCCAATCTGACTCTCCCGC 1258
Db 1213 GACTTTAAATGATACACCACTTACCTGTGAGGGGAGCCCAATCTGACTCTCCCGC 1272
Qy 1259 CTGAGCACTGCAAGCGCGGAGAGAGTCCCGCCAGGCTGTGGCCAGAGAGCTTCAG 1318
Db 1273 CTGAGCACTGCAAGCGCGGAGAGAGTCCCGCCAGGCTGTGGCCAGAGAGCTTCAG 1332
Qy 1319 AAGGGCACTGAGCGCTGCGGCGAGGCTCGAGCATCGAGAGGCAACGAGGAAAGTCT 1378
Db 1333 AAGGGCACTGAGCGCTGCGGCGAGGCTCGAGCATCGAGAGGCAACGAGGAAAGTCT 1392
Qy 1379 CCGGGGCGATCTGTAATTAACCTTTTCTTTTGT 1421
Db 1393 CCGGGGCGATCTGTAATTAACCTTTTCTTTTGT 1435

RESULT 9
ABZ67235
ID ABZ67235 standard; cDNA; 1460 BP.
XX
AC ABZ67235;
XX
DT 26-MAR-2003 (first entry)
DE
XX Human secreted protein encoding cDNA SEQ ID NO 355.
XX
KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
KW vulnerary; antibacterial; antiparkinsonian; antischizoid; antianemic;
KW antiallergic; cancer; antihemetic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antilucer; anticonvulsant;
KW antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephroretropic;
KW gene therapy; gene; chromosome 9p21; ds.
XX
OS Homo sapiens.
XX
PN MO20027186-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002MO-US009188.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2003-040583/03.
DR P-PSDB; ABP99814.
XX
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX
PS Claim 7; Page 1353; 2423pp; English.
XX
XX The invention relates to novel human genes (ABZ6891-ABZ68209) and the
CC encoded secreted proteins (ABP9470-ABP9872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and

CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections
 CC XX

SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match 96.8%; Score 1375.8; DB 10; Length 1460;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

QY 1 GAACAGAAAGTAAAGGCTTACAGACAGTGAATAATGTTCCGTCGCGGCTAGAAAAA 60
 DB 14 GAACAGAAAGTAAAGGCTTACAGACAGTGAATAATGTTCCGTCGCGGCTAGAAAAA 73
 QY 61 CTCTGTCCGTAACCAACCCCAAGCGTTGAGAGACGCCACCTTCACGCTTCTTAAAGGA 120
 DB 74 CTCTGTCCGTAACCAACCCCAAGCGTTGAGAGACGCCACCTTCACGCTTCTTAAAGGA 133
 QY 121 GAGGTGACAGACTCAAGACTTCAACAGCCCACTGGTCCCAAGCCTTGTAAAGCAAGAGAG 180
 DB 134 GAGGTGACAGACTCAAGACTTCAACAGCCCACTGGTCCCAAGCCTTGTAAAGCAAGAGAG 193
 QY 181 CCAAGAGCGGCTCTCCCGGCTCAGGACAGCCCAAGCTTGTGCTGCTGCGCGCTG 240
 DB 194 TCAAGAGCGGCTCTCCCGGCTCAGGACAGCCCAAGCTTGTGCTGCTGCGCGCTG 253
 QY 241 CGTGACAGACTCGGCGGCTGACAGATGACCTGTGAAACGCGGTAAGTCTGCTTTTAA 300
 DB 254 CGTGACAGACTCGGCGGCTGACAGATGACCTGTGAAACGCGGTAAGTCTGCTTTTAA 313
 QY 301 CCGAGCCCGGCAATGCGGAGGCTTCAAGCGTTCACGCTCAGCTGATCTGATGTTT 360
 DB 314 CCGAGCCCGGCAATGCGGAGGCTTCAAGCGTTCAGCTCAGCTGATCTGATGTTT 373
 QY 361 TGGCTCAGAGCAAGCTTCTGCTCATCTTCCGCGGATCCGAGGCACTCGCGCTGCT 420
 DB 374 TGGCTCAGAGCAAGCTTCTGCTCATCTTCCGCGGATCCGAGGCACTCGCGCTGCT 433
 QY 421 TTTGGTTGTAAGATCTTCTCAGTCTGTTCAATGAGCGCAAAATTTGGCTGTGCACT 480
 DB 434 TTTGGTTGTAAGATCTTCTCAGTCTGTTCAATGAGCGCAAAATTTGGCTGTGCACT 493
 QY 481 TCAATGCAATGATGTTCTGTGGGTACAGTGAACACCAACATCTTCAAAAGCTTCAAG 540
 DB 494 TCAATGCAATGATGTTCTGTGGGTACAGTGAACACCAACATCTTCAAAAGCTTCAAG 553
 QY 541 CAGGCGCGTTAAGCGCGGTGTGCTGCTCGGCGCTGAGAGGCAATTAATTAAC 600
 DB 554 CAGGCGCGTTAAGCGCGGTGTGCTGCTCGGCGCTGAGAGGCAATTAATTAAC 613
 QY 601 TCAAGGAGACCCCAAGTCAATGATGTAACAGACATTTGATCAACAGAGAGTTCACCT 660
 DB 614 TCAAGGAGACCCCAAGTCAATGATGTAACAGACATTTGATCAACAGAGAGTTCACCT 673
 QY 661 GGGCTTGAAGAGATTAAGCGCGGAGTACGGAACGCACTGTGAAGAAAGGGGCTGCGCG 720
 DB 674 GGGCTTGAAGAGATTAAGCGCGGAGTACGGAACGCACTGTGAAGAAAGGGGCTGCGCG 733
 QY 721 ACCCAGTCTTACCTGCGGAGAGATTAACAGAGAGGCTTGGCGCTGTACACAC 780
 DB 734 ACCCAGTCTTACCTGCGGAGAGATTAACAGAGAGGCTTGGCGCTGTACACAC 793
 QY 781 AGTACCACTGGCGGAGCACTACGCTTGGCGACGCTATGAGTGGCTTCTGCTTCTGGC 840
 DB 794 AGTACCACTGGCGGAGCACTACGCTTGGCGACGCTATGAGTGGCTTCTGCTTCTGGC 853

QY 841 TCTCTCAACAGTGTCTCTCCACAGCGCGGCCCGCTTACAGGAGCGCTGACACTGTA 900
 DB 854 TCTCTCAACAGTGTCTCTCCACAGCGCGGCCCGCTTACAGGAGCGCTGACACTGTA 913
 QY 901 CCAACGAGCCTTCCGCTCTTCCGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 960
 DB 914 CCAACGAGCCTTCCGCTCTTCCGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 973
 QY 961 TCTGCGCGCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1020
 DB 974 TCTGCGCGCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1033
 QY 1021 TCAAGCTGCAACCGGCGCTCTGCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1080
 DB 1034 TCAAGCTGCAACCGGCGCTCTGCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1093
 QY 1081 ATGTTCCGCGCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1139
 DB 1094 ATGTTCCGCGCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1153
 QY 1140 GAGAGAGGCGCTCACTCTTATCCCTGCGGACCCCACTGCAAGAGAGCGCTCCCA 1199
 DB 1154 AAGAGGCGGCGCTCACTCTTATCCCTGCGGACCCCACTGCAAGAGAGCGCTCCCA 1212
 QY 1200 GAC-TTAAATGATCACTAATCTGAGAGGAGACCAATCTGACTCTTCCCGC 1258
 DB 1213 GACTTTAAATGATCACTAATCTGAGAGGAGACCAATCTGACTCTTCCCGC 1272
 QY 1259 CTTGGACATCGAGGCGCGGAAAGAGTGCAGCGCTGCGGCGGAGAGTCTCAGG 1318
 DB 1273 CTTGGACATCGAGGCGCGGAAAGAGTGCAGCGCTGCGGCGGAGAGTCTCAGG 1332
 QY 1319 AAGGGACATGAGGCGCTGCGGCGGAGAGTGCAGCGCTGCGGCGGAGAGTCT 1378
 DB 1333 AAGGGACATGAGGCGCTGCGGCGGAGAGTGCAGCGCTGCGGCGGAGAGTCT 1392
 QY 1379 CTTGGGCGATCTGTAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1421
 DB 1393 CTTGGGCGATCTGTAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1435

RESULT 10
 ADU01671
 ID ADU01671 standard; cDNA; 960 BP.
 AC ADU01671;
 AC ADU01671;
 DT 27-JAN-2005 (first entry)
 XX
 DE Novel human polynucleotide seqid 138.
 XX cytoabatic; antiproliferic; antiinflammatory; gene therapy; Nanodisc;
 KW proliferative disorder; inflammatory disorder; immune disorder;
 KW metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
 KW ulcerative colitis; human; gene; ss.
 XX
 OS Homo sapiens.
 PN
 XX WO2004093804-A2.
 PD 04-NOV-2004.
 PD 04-NOV-2004.
 XX
 PF 19-APR-2004; 2004WO-US012047.
 XX
 XX 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 PR 08-SEP-2003; 2003US-0505059P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Lee E, Heestir K, Chu K, Masuoka L, Williams LT;
 PI MPI; 2004-775861/76.
 DR P-PSDB; ADU02403.

XX New first nucleic acid molecule comprising a polynucleotide sequence
 PT given in the specification, useful in preparing a composition for
 PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.

PS Claim 1, SEQ ID NO 138; 291pp; English.

CC The invention describes a new first nucleic acid molecule comprising a
 CC polynucleotide sequence given in the specification. Also described are:
 CC an animal injected with the nucleic acid molecule; a second nucleic acid
 CC molecule comprising a second polynucleotide sequence that is at least
 CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
 CC that hybridises to the first polynucleotide sequence under high
 CC stringency conditions; a vector comprising the nucleic acid molecule and
 CC a promoter that drives the expression of the nucleic acid molecule; a
 CC host cell transfected, transduced or infected with the
 CC nucleic acid molecule; a nucleic acid composition comprising the nucleic acid
 CC a buffer and one or more compositions comprising the nucleic acid
 CC molecule, vector or host cell; a substantially purified polypeptide; an
 CC animal injected with the polypeptide; a polypeptide composition
 CC comprising the polypeptide molecule and a carrier or buffer; a cell
 CC culture medium comprising the polypeptide or transfected cells
 CC transfected with the polynucleotide; making a transformed, transduced,
 CC transduced, or infected host cell; synthesising Nanodiscs simultaneously
 CC and for synthesising a series of simultaneously-synthesised Nanodiscs
 CC sequentially utilising a dynamic system; preparing a hydrophobic protein
 CC for determination of crystal structure; immunising a non-human animal;
 CC screening for modulators of hydrophobic protein activity; a diagnostic
 CC kit; determining the presence of the nucleic acid molecule or its
 CC complement; determining the presence of an antibody to the polypeptide in
 CC a sample; an antibody specifically recognising, binding to or modulating
 CC the biological activity of at least one polypeptide encoded by a nucleic
 CC acid molecule or its biologically active fragment; an antibody
 CC composition comprising the antibody and a carrier; a bacteriophage, where
 CC the antibody is displayed on the bacteriophage; a bacterial cell
 CC comprising the bacteriophage; a non-human animal injected with the
 CC antibody composition; a host cell that secretes the antibody; making an
 CC antibody; diagnosing a disease, disorder, syndrome, or condition
 CC comprising cancer, or proliferative, inflammatory, immune, syndromes or
 CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
 CC conditions in a patient; a modulator composition comprising a modulator
 CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
 CC subject; an isolated modified cell comprising at least one first
 CC heterologous nucleic acid molecule, where the first heterologous nucleic
 CC acid molecule comprises a first polynucleotide sequence that encodes a
 CC first polypeptide; a non-human animal deficient in the polypeptide or
 CC that over-expresses the polypeptide; isolated tissues derived from the
 CC non-human animal; and one or more cells derived from the non-human
 CC animal. The nucleic acid is useful in preparing a composition for
 CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
 CC This sequence encodes a novel human polypeptide of the invention.

XX Sequence 960 BP, 168 A, 321 C, 260 G, 211 T, 0 U, 0 Other;

QY Query Match 67.4%; Score 958.4; DB 13; Length 960;
 DB Best Local Similarity 99.9%; Pred. No. 1.8e-229;
 Matches 959; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 267 ATGACCTGTGGAAAGCGGCTCTGCTTTTAAACCCCAAGCCCGGAGTCCGAGGCTTC 326
 DB 1 ATGACCTGTGGAAAGCGGCTCTGCTTTTAAACCCCAAGCCCGGAGTCCGAGGCTTC 60

QY 327 AGCGTTCACATGCTCATCGTTATTTCTAGTGTGTTTGGCTCTAGACAGCAAGCTTCCTGTC 386
 DB 61 AGCGTTCACATGCTCATCGTTATTTCTAGTGTGTTTGGCTCTAGACAGCAAGCTTCCTGTC 120
 QY 387 ATCTTGGCGGGAGATCCGTGGCCACTCGCGCTGTGTTTGGTTGGTAGAGTTCTTCTCAGT 446
 DB 121 ATCTTGGCGGGAGATCCGTGGCCACTCGCGCTGTGTTTGGTTGGTAGAGTTCTTCTCAGT 180
 QY 447 CTGTTTCATAGGCGCAAGAAATTGTGCTGTGCACTTCAGTGCAGAAATGTTGGTGGTAC 506
 DB 181 CTGTTTCATAGGCGCAAGAAATTGTGCTGTGCACTTCAGTGCAGAAATGTTGGTGGTAC 240
 QY 507 GTGAACACCAACATCTCTACAAAGCCTTCAGCGAGCGCGGTTACAGCCCGTCTCGT 566
 DB 241 GTGAACACCAACATCTCTACAAAGCCTTCAGCGAGCGCGGTTACAGCCCGTCTCGT 300
 QY 567 CTGCTGTGGGCTGTGAGAGGCTTAAATTATACCTCAAGAGGAGCCCAAGTGCATAGCGT 626
 DB 301 CTGCTGTGGGCTGTGAGAGGCTTAAATTATACCTCAAGAGGAGCCCAAGTGCATAGCGT 360
 QY 627 AACGAGACCATTAATGATACAAAGCAAGTTCACCTGCGCTTGAAGAAATTAACGCGCG 686
 DB 361 AACGAGACCATTAATGATACAAAGCAAGTTCACCTGCGCTTGAAGAAATTAACGCGCG 420
 QY 687 GAGTACGCGAAGCACTGAGAGAGGGCTGCCGAAACCAATGCTCTTACTGCGGAGAG 746
 DB 421 GAGTACGCGAAGCACTGAGAGAGGGCTGCCGAAACCAATGCTCTTACTGCGGAGAG 480
 QY 747 TTTCACACGAGTAAAGCTTTGCGGCTGTACCAACAGTACCACTGCGGAGCACTACAGCGC 806
 DB 481 TTTCACACGAGTAAAGCTTTGCGGCTGTACCAACAGTACCACTGCGGAGCACTACAGCGC 540
 QY 807 TCGGCGACGCTAATGAGGTGGCGTTTGTGCTTCTGCTCTTCCAAAGTGTGCTTCCACG 866
 DB 541 TCGGCGACGCTAATGAGGTGGCGTTTGTGCTTCTGCTCTTCCAAAGTGTGCTTCCACG 600
 QY 867 CCGGCGCCGCTCTACGAGAGGCTGTGACCTGTCACCAACCGAGCCTTCGCTCTTCGCG 926
 DB 601 CCGGCGCCGCTCTACGAGAGGCTGTGACCTGTCACCAACCGAGCCTTCGCTCTTCGCG 660
 QY 927 GTCTTGGCTTTGGCTTCATCTTACGCGCGCTGCGCGCTCGGCTACGAGCTCTCC 986
 DB 661 GTCTTGGCTTTGGCTTCATCTTACGCGCGCTGCGCGCTCGGCTACGAGCTCTCC 720
 QY 987 GCGCTACCACTCAAGTACGAGCGCGCTTCTGAGTACCGCTGCAACCGAGCTCTGTGC 1046
 DB 721 GCGCTACCACTCAAGTACGAGCGCGCTTCTGAGTACCGCTGCAACCGAGCTCTGTGC 780
 QY 1047 CTCTTCTCGAGAGGGCGGTGTAGTCTTCAGATGTTTCGGCCGACGCTCTTGCGACG 1106
 DB 781 CTCTTCTCGAGAGGGCGGTGTAGTCTTCAGATGTTTCGGCCGACGCTCTTGCGACG 840
 QY 1107 CTCTTGGACCAAAAGCGCCAAAGACTGCGACGCAAGAGAGGGGGCTCACCTTATCTTC 1166
 DB 841 CTCTTGGACCAAAAGCGCCAAAGACTGCGACGCAAGAGAGGGGGCTCACCTTATCTTC 900
 QY 1167 GCGGACCACTGCAAGAGAGCGCTCTCCAGACTTAAATGTATCAACCACTAACCTG 1226
 DB 901 GCGGACCACTGCAAGAGAGCGCTCTCCAGACTTAAATGTATCAACCACTAACCTG 960

RESULT 11
 ACF58262
 ID ACF58262 standard; DNA; 981 BP.
 AC ACF58262;
 AC 12-FEB-2004 (first entry)
 DT Human chr15:41,013.a nucleotide sequence.
 XX Colon cancer; cytostatic; gene therapy; colorectal tumour; human; gene;
 KM de.

protein, methods and compositions have cytostatic and immunostimulant activities. The nucleic acids may be used to treat disorders by gene therapy. The nucleic acids and proteins, methods and compositions are useful in diagnosing and treating cancer, e.g. cancer of the colon, pancreas, breast, ovary or lung, and in inducing an immune response. This polynucleotide represents a DNA sequence relating to the overexpressed nucleic acids of human cancer cells of the invention.

Sequence 981 BP; 171 A; 315 C; 294 G; 201 T; 0 U; 0 Other;

Query Match 54.1%; Score 768.8; DB 12; Length 981;
Best Local Similarity 99.7%; Pred. No. 4,9e-182;
Matches 770; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 267 ATGACCCCTGTGAAGCGGCTACTGCTTTTAAACCCCGGCGCATGCCGAGGCTTC 326
DB 1 ATGACCCCTGTGAAGCGGCTACTGCTTTTAAACCCCGGCGCATGCCGAGGCTTC 60
QY 327 AGCGTTCATGCTCATGCTATTTCTAGTGTGTTTGGCTTGAAGCAAGCTTCGCTC 386
DB 61 AGCGTTCATGCTCATGCTATTTCTAGTGTGTTTGGCTTGAAGCAAGCTTCGCTC 120
QY 387 ATCTTGGCGGGGATCCGCGGCACTCGCGCTGTTTGGTTGGTGAAGTTCCTTCAGT 446
DB 121 ATCTTGGCGGGATCCGCGGCACTCGCGCTGTTTGGTTGGTGAAGTTCCTTCAGT 180
QY 447 CTGTTCATGAGCGGAGAAATGTGCTGTGTGCACTTCAGTGAGAAATGTTGCGGGTACA 506
DB 181 CTGTTCATGAGCGGAGAAATGTGCTGTGTGCACTTCAGTGAGAAATGTTGCGGGTACA 240
QY 507 GTGAACACCAACATCTTAAGAAAGCTTCAAGCGGAGCGGCTTACAGCCGCTGTGCT 566
DB 241 GTGAACACCAACATCTTAAGAAAGCTTCAAGCGGAGCGGCTTACAGCCGCTGTGCT 300
QY 567 CTGCTGTGGGCTGTGAAGGCACTTAATATTAACAACAAGGACCCAGTGCATCAGCTG 626
DB 301 CTGCTGTGGGCTGTGAAGGCACTTAATATTAACAACAAGGACCCAGTGCATCAGCTG 360
QY 627 AACGAGACATTAAGCTTAACAAGAGAGTTCACCTGGCTGTGAAGAAATTAACGGCGG 686
DB 361 AACGAGACATTAAGCTTAACAAGAGAGTTCACCTGGCTGTGAAGAAATTAACGGCGG 420
QY 687 GAGTAGCGGAAGCGACGAGAGAGGCGTCCGACCCAGTGTCTTAACCTGTGCGGAGAA 746
DB 421 GAGTAGCGGAAGCGACGAGAGAGGCGTCCGACCCAGTGTCTTAACCTGTGCGGAGAA 480
QY 747 TTCAACCGAGTAGCCCTTGGCGCTGTACACCAAGTACCACTGGCGGAGCACTAGCGC 806
DB 481 TTCAACCGAGTAGCCCTTGGCGCTGTACACCAAGTACCACTGGCGGAGCACTAGCGC 540
QY 807 TCGGCGACGCTAAGGAGGCGTCTGTGCTTGTGCTCTCTCCAAAGTGTGCTTTCACG 866
DB 541 TCGGCGACGCTAAGGAGGCGTCTGTGCTTGTGCTCTCTCCAAAGTGTGCTTTCACG 600
QY 867 CCGGCGCCGCTTACGAGAGGCTGTGCACTGTGACCAACCGAGGCTTTCGCGCTTCGGG 926
DB 601 CCGGCGCCGCTTACGAGAGGCTGTGCACTGTGACCAACCGAGGCTTTCGCGCTTCGGG 660
QY 927 GTCTTGGCTTGGCTTTCATCTTAAGGTCGCTGTGCGCTGTGCGCTTGAAGCTTCCTC 986
DB 661 GTCTTGGCTTGGCTTTCATCTTAAGGTCGCTGTGCGCTGTGCGCTTGAAGCTTCCTC 720
QY 987 GCGCTACCACTAGTAGCGGCGCGCTTCTGGGTTCAGCGTGAAGCAACCGGCG 1038
DB 721 GCGCTACCACTAGTAGCGGCGCGCTTCTGGGTTCAGCGTGAAGCAACCGGCG 772

```

RESULT 13

AAA94624
ID AAA94624 standard; DNA; 498 BP.
XX
AC AAA94624;
XX

```

DT 11-JAN-2001 (first entry)
XX Human CASB618 EST.
DE
XX
XX
XX Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW colon; autoimmune disease; HLA_A0201; expressed sequence tag; EST; ss.
XX
XX Homo sapiens.
XX
XX WO200053748-A2.
XX
XX 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-EP002048.
XX
XX 11-MAR-1999; 99GB-00005607.
XX
XX 01-SEP-1999; 99GB-00020590.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Bruck CEM, Cassart J, Coche T, Vinals Y De Basolac,
PI WPI, 2000-572268/53.
XX
XX New human CASB618 polypeptide, useful as a vaccine for prophylactic and
PT therapeutic treatment of cancers, particularly ovarian or colon cancer,
PT autoimmune diseases and related conditions.
XX
XX Claim 32; Page 62; 76pp; English.
XX
XX
XX The present sequence is an expressed sequence tag (EST) for human CASB618
CC protein. The gene for human CASB618 is thought to be located on
CC chromosome 15. CASB618 protein and epitopes of CASB618 protein (see
CC AAS26327 to AAS26399) are useful in diagnosing the occurrence of tumour
CC cells and in vaccines for prophylactic and therapeutic treatment of
CC cancers, particularly ovarian or colon cancer, autoimmune diseases and
CC related conditions
XX
XX
XX Sequence 498 BP; 97 A; 166 C; 137 G; 98 T; 0 U; 0 Other;
SO

```

Query Match 31.9%; Score 453; DB 3; Length 498;
Best Local Similarity 99.6%; Pred. No. 4e-103;
Matches 475; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

QY 947 CTCTAGGCGTGGCGCTTGGCGGCTTCTGCGCTTCTCTGCGAGGAGCC 1005
DB 1 CTCTAGGCGTGGCGCTTGGCGGCTTCTGCGCTTCTCTGCGAGGAGCC 60
QY 1006 -GCGCGGCTTCTGAGTCAAGCTGGAACCGGCGTCTGTGCTTCTCTGAGGAGCC 1064
DB 61 AGCGCGGCTTCTGAGTCAAGCTGGAACCGGCGTCTGTGCTTCTCTGAGGAGCC 120
QY 1065 GTGATGAGTCTCCAGTATGTTTGGCCGAGGCTTGTGCAACCTTCTGACCAAGGCGC 1124
DB 121 GTGATGAGTCTCCAGTATGTTTGGCCGAGGCTTGTGCAACCTTCTGACCAAGGCGC 180
QY 1125 AAGGACTGCAAGCAGAGAGGAGGAGGCTTCACTTATCTTGGCGAGCCACGACGACAAG 1184
DB 181 AAGGACTGCAAGCAGAGAGGAGGAGGCTTCACTTATCTTGGCGAGCCACGACGACAAG 240
QY 1185 CAGGCGCTTCTCCAGACTTAATAATGTATACCACTTAACCTGTGAAGGAGGCCAAATCTG 1244
DB 241 CAGGCGCTTCTCCAGACTTAATAATGTATACCACTTAACCTGTGAAGGAGGCCAAATCTG 300
QY 1245 GACTCTTCTCCGCGCTTGGGAGATTCGAGGCGGAGAGGAGGCGGCGGAGGCTTGGGCGC 1304
DB 301 GACTCTTCTCCGCGCTTGGGAGATTCGAGGCGGAGAGGAGGCGGCGGAGGCTTGGGCGC 360
QY 1305 AGGAGAGCTTCAAGAAAGGCACTGAAGCGCTGTGCGGAGGAGGCTTCAAGAGGAGC 1364
DB 361 AGGAGAGCTTCAAGAAAGGCACTGAAGCGCTGTGCGGAGGAGGCTTCAAGAGGAGC 420
QY 1365 ACCAGGAAAGTCTCTGAGGAGGAGTCTGTAATAAAGCTTTTCTTTTGTGTTTTT 1421

```

Db 421 ACCAGGAAAGTCTCTGAGGCGATCTGTAATAACCTTTTCTTTGTTTTT 477

RESULT 14

ID AB274587 standard; DNA; 5033 BP.

AC AB274587;

DT 12-MAY-2003 (first entry)

DE Secreted protein gene 360 genomic fragment HUFCL31, SEQ ID NO:1734.

XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KM acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KM drug screening; chromosome identification; chromosome mapping;
 KM cytotoxic; gene therapy; anti-inflammatory; immunomodulator; anti-HIV;
 KM antineoplastic; vulnery; chromosome 9p21; gene; ds.

OS Homo sapiens.

PN MO200277013-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002MO-US009370.

PR 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-040578/03.

PT New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.

PS Disclosure; Page 2315-2316; 2474bp; English.

XX AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC AB273698-AB274687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prothormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention

XX Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

XX Query March 29.1%; Score 413.4; DB 8; Length 5033;

XX Best Local Similarity 98.6%; Pred. No. 6.7e-93;

XX Matches 417; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAACGAGAAAGTAAAGGCTACAGACAGTGAAGAAATAGTTTGTCTGCGCGCTAGAAAA 60

Db 5 GAACGAGAAAGTAAAGGCTACAGACAGTGAAGAAATAGTTTGTCTGCGCGCTAGAAAA 64

QY 61 CTCTGTGGTACCAACCCGAGGCGTTGAGACAGCCCACTTCCGCTTAAAGGA 120

Db 65 CTCTGTGGTACCAACCCGAGGCGTTGAGACAGCCCACTTCCGCTTAAAGGA 124

QY 121 GAGGTGACAGATCTGACACTTTCACAGCCCACTGCGCTTGTAGCAAAAGAGAG 180

Db 125 GAGGTGACAGATCTGACACTTTCACAGCCCACTGCGCTTGTAGCAAAAGAGAG 184

QY 181 CCAAGGAGCGGCTCTCCGCGTTCAGAGGAGCCCACTTGTGCTTGTGCGCGCTG 240

Db 185 CCAAGGAGCGGCTCTCTCCGCGTTCAGAGGAGCCCACTTGTGCTTGTGCGCGCTG 244

QY 241 CGTGACAGACTGCGCGCGGTGACAGATGACCTGTGAAAGGCGTACTGCTTTTACC 300

Db 245 CGTGACAGACTGCGCGCGGTGACAGATGACCTGTGAAAGGCGTACTGCTTTTACC 304

QY 301 CCCAGCCCCGCGATGCGCGAGGCTTCAAGCTTCACTGCTCATTCCTTATTAAGTCTTT 360

Db 305 CCCAGCCCCGCGATGCGCGAGGCTTCAAGCTTCACTGCTCATTCCTTATTAAGTCTTT 364

QY 361 TGGCTTACAGACCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGACACTGCGCTGGT 420

Db 365 TGGCTTACAGACCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGACACTGCGCTGGT 424

QY 421 TTT 423

Db 425 TGT 427

RESULT 15

ID ADC21005 standard; DNA; 5033 BP.

AC ADC21005;

DT 18-DEC-2003 (first entry)

DE Human secreted protein-related DNA sequence #423.

XX gene therapy; human; secreted protein; haemopoietic disorder;
 KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
 KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
 KW leukaemia; wound healing; epithelial cell proliferation disorder;
 KW immune disorder; autoimmune disorder; asthmatic disorder;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
 KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

OS Homo sapiens.

PN MO200292787-A2.

PD 21-NOV-2002.

PF 26-MAR-2002; 2002MO-US009257.

PR 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-129287/12.

PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,

PT preventing or treating hematopoietic or hematologic disorders, e.g.
XX anemia or hemophilia.

PS Disclosure, SEQ ID NO 959; 1512pp; English.

XX
CC The invention comprises the amino acid and coding sequences of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful for detecting, preventing, diagnosing, prognosticating, treating
CC or ameliorating; hematopoietic or hematological disorders (e.g. anaemia
CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC and Crohn's disease); neoplastic disease (e.g. cancer and leukemia);
CC wound healing and disorders of epithelial cell proliferation; immune
CC disorders (e.g. autoimmune disorders and asthmatic disorders);
CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and
CC gastroenteritis). The present DNA sequence was used in the
CC exemplification of the invention.

SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 29.1%; Score 413.4; DB 10; Length 5033;

Best Local Similarity 98.6%; Pred. No. 6,7e-93;

Matches 417; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	GAACGAGAAAGTACGCGCTACAGACAGTGAATAATGTTGCTGCGCGCTAAGAAAA	60
DB	5	GAACGAGAAAGTACGCGCTACAGACAGTGAATAATGTTGCTGCGCGCTAAGAAAA	64
QY	61	CTTGTGCTGCTACCAACCCGAGCGTTAGAGCAAGCCACCTCCACGCTTCTTAAAGGA	120
DB	65	CTTGTGCTGCTACCAACCCGAGCGTTAGAGCAAGCCACCTCCACGCTTCTTAAAGGA	124
QY	121	GAGGTGAGGACTCAGACTTACACAGCCACTGGTCCAGCCTGTGTAAGCAAGAGACG	180
DB	125	GAGGTGAGGACTCAGACTTACACAGCCACTGGTCCAGCCTGTGTAAGCAAGAGACG	184
QY	181	CCAAGGACGGGCTCTCCCGGCTCAGGCAAGCCCACTTGCTTGCTTGCCGCTG	240
DB	185	CCAAGGACGGGCTCTCCCGGCTCAGGCAAGCCCACTTGCTTGCTTGCCGCTG	244
QY	241	CGTGCAGCACTCGCGCGGCGCTGCAAGTACCTGTGAAAGGCGTACTGCTTTTACC	300
DB	245	CGTGCAGCACTCGCGCGGCGCTGCAAGTACCTGTGAAAGGCGTACTGCTTTTACC	304
QY	301	CCGAGCCCGGCGATGCGCGAGGCTTCAAGGCTTCCATGCTCATGTTTCTAGTGT	360
DB	305	CCGAGCCCGGCGATGCGCGAGGCTTCAAGGCTTCCATGCTCATGTTTCTAGTGT	364
QY	361	TGGCTTAGAGCAAGCTTCTGCTCATCTTGGCCGGGATCCGTGACCACTCGCGTGGT	420
DB	365	TGGCTTAGAGCAAGCTTCTGCTCATCTTGGCCGGGATCCGTGACCACTCGCGTGGG	424
QY	421	TTT 423	
DB	425	TGT 427	

Search completed: January 21, 2006, 22:44:27
Job time : 940 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:54:27 ; Search time 6251 Seconds
(without alignments)
10635.809 Million cell updates/sec

Title: US-09-937-059-57

Perfect score: 1421
Sequence: 1 gaaccaggaagtaacgcgt.....ctttttctttgttttt 1421

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gsa1.*
10: gb_gsa2.*
11: gb_gsa3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	963	67.8	963	11	DQ050095 Homo sapi
2	701.6	49.4	769	11	DQ050096 Pan trogl
3	693	48.8	1594	4	AK018569 Mus muscu
4	657	46.2	1309	4	AK008816 Mus muscu
5	615.8	43.3	619	8	DN831248 MGCL.5.1.
6	610.6	43.0	617	8	DN831247 MGCL.5.1.
7	602	42.4	602	8	DN930503 MGCL.5.1.
8	600.8	42.3	622	8	DN930502 MGCL.5.1.
9	529.6	37.3	537	1	AI821606 nk08a11.x
10	526.4	37.0	717	7	CK971974 4087846.B
11	522.2	36.7	703	7	CK979918 4111631.B
12	512	36.0	515	1	AI826629 wk35e04.x
13	510	35.9	510	1	AI304327 q057h06.x
14	496	34.9	500	3	BP396416 BP396416
15	489.2	34.4	505	1	AI791844 nk08a11.Y
16	468.8	33.0	704	7	CK963936 4078848.B
17	468.2	32.9	886	7	CU022684 CU022684
18	467.4	32.9	886	7	AI660560 we68b08.x
19	466	32.8	1014	5	BY708714 BY708714
20	456.8	32.1	716	7	CK981626 4114147.B
21	450	31.7	849	7	CU023383 CU023383
22	445.4	31.3	466	1	AA573825 nk08a11.S

C	23	443	31.2	446	1	AI984141	AI984141 wu1c02.x
C	24	433	30.5	434	1	AI991272	AI991272 wu1h04.x
C	25	432.2	30.4	826	3	BP435528	BP435528 BP435528
C	26	429	30.2	595	8	DN350567	DN350567 LB3579-0
C	27	418	29.4	876	3	BP440241	BP440241 BP440241
C	28	414	29.1	414	1	AA593860	AA593860 nm19f03.8
C	29	405	28.5	405	1	AI274929	AI274929 q149c11.x
C	30	404.4	28.5	411	1	AI983793	AI983793 wu20c09.x
C	31	400.2	28.2	405	1	AI346155	AI346155 qp43f12.x
C	32	391	27.5	538	1	AI660493	AI660493 we67h02.x
C	33	389.8	27.4	806	7	CK789912	CK789912 ACBNCOURT
C	34	388.4	27.3	538	5	BX517021	BX517021 BX517021
C	35	381	26.8	381	1	AI281211	AI281211 qk58e08.x
C	36	380.4	26.8	747	7	CU027584	CU027584 CU027584
C	37	380	26.7	573	6	CB854121	CB854121 UI-CF-DU1
C	38	375.8	26.4	737	6	CB305399	CB305399 UI-CF-BN1
C	39	373.6	26.3	660	6	CF766127	CF766127 CBS004870
C	40	372.8	26.2	646	5	BY720747	BY720747 BY720747
C	41	360.4	25.4	713	6	CB171450	CB171450 JID602700
C	42	356.4	25.1	1032	10	AY418362	AY418362 Homo sapi
C	43	354.6	25.0	698	7	CK968084	CK968084 4083252.B
C	44	352.2	24.8	485	3	BP433235	BP433235 BP433235
C	45	350.4	24.7	671	7	CN789601	CN789601 4124079.B

ALIGNMENTS

RESULT 1	DQ050095	963 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ050095				
DEFINITION	Homo sapiens HC18097 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	DQ050095				
VERSION	DQ050095.1	GI:69303294			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Friedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Smitsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees				
JOURNAL	(et) Plos Biol. 3 (6), E170 (2005)				
PUBMED	15869325				
REFERENCE	2 (bases 1 to 963)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Friedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Smitsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission.				
JOURNAL	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering this based on alignment. Translation starts at the beginning of alignment.				
FEATURES	Location/Qualifiers				
SOURCE	1..963				
gene	/organism="Homo sapiens"				
ORIGIN	/mol_type="genomic DNA"				
Query Match	67.8%; Score 963; DB 11; Length 963;				
Best Local Similarity	100.0%; Pred. No. 6,3e-236; Indels 0;				
Matches	963; Conservative 0; Mismatches 0; Gaps 0;				

```

QY 267 ATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCCGGAGTCCGAGGCTTC 326
Db 1 ATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCCGGAGTCCGAGGCTTC 60
QY 327 AGCGTTCACTGTGCTACATCGTTATTTAGTGTGTTTGGCTCTAGACAGAGCTTCGTCTC 386
Db 61 AGCGTTCACTGTGCTACATCGTTATTTAGTGTGTTTGGCTCTAGACAGAGCTTCGTCTC 120
QY 387 ATCTTGGCCGGGGATCCGTGGCCCATCGCGCTGGTGTGTTGGTGTGAGTCTTCTCACT 446
Db 121 ATCTTGGCCGGGGATCCGTGGCCCATCGCGCTGGTGTGTTGGTGTGAGTCTTCTCACT 180
QY 447 CTGTTATAGGCGCAGAAATTTGGCTGTGCACTTCACTGACAGAAATGTTTCGTGGTACA 506
Db 181 CTGTTATAGGCGCAGAAATTTGGCTGTGCACTTCACTGACAGAAATGTTTCGTGGTACA 240
QY 507 GTGAACACCAACATCTTCAAAAGCCTTCAAGCGCAGCGCGCTTACAGCCGTCGTGGT 566
Db 241 GTGAACACCAACATCTTCAAAAGCCTTCAAGCGCAGCGCGCTTACAGCCGTCGTGGT 300
QY 567 CTGCTGTGGGCGCTGAGGGCATTATATTTATACCTGACAGGGACCCCACTGATCAGCTG 626
Db 301 CTGCTGTGGGCGCTGAGGGCATTATATTTATACCTGACAGGGACCCCACTGATCAGCTG 360
QY 627 AACGAGACCATTTGACTACAAAGAGAGTTCACCTGGCGTCTGAAGAAGATTACGCGCG 686
Db 361 AACGAGACCATTTGACTACAAAGAGAGTTCACCTGGCGTCTGAAGAAGATTACGCGCG 420
QY 687 GAGTACGGAACGCACTGAGAAAGGGGCTGCGGACCCAGTGTCTTACCTGGCGGAGAA 746
Db 421 GAGTACGGAACGCACTGAGAAAGGGGCTGCGGACCCAGTGTCTTACCTGGCGGAGAA 480
QY 747 TTCAACACGAGTAGGCTTGGCGGCTGTACCAACAGTACCACTGGCGGAGACCTACAGCC 806
Db 481 TTCAACACGAGTAGGCTTGGCGGCTGTACCAACAGTACCACTGGCGGAGACCTACAGCC 540
QY 807 TCGGCAACGCTATAGGTGGTGTGCTGTGCTCTTCCAAAGTGTGCTCTTCACAG 866
Db 541 TCGGCAACGCTATAGGTGGTGTGCTGTGCTCTTCCAAAGTGTGCTCTTCACAG 600
QY 867 CCGGCCCCGCTCTACGGAAGGCTGTGCACTGTGCAACCGGAGCCTTCGCGCTCTTCGCG 926
Db 601 CCGGCCCCGCTCTACGGAAGGCTGTGCACTGTGCAACCGGAGCCTTCGCGCTCTTCGCG 660
QY 927 GTCTTGCCTTGGCTCTCACTCTAGAGGCGCTGTGCGCGCTCCGCTCCAGGCTCTCC 986
Db 661 GTCTTGCCTTGGCTCTCACTCTAGAGGCGCTGTGCGCGCTCCGCTCCAGGCTCTCC 720
QY 987 GCGCTCACCACTAGTACGCGCGCGCTTCTGGGTACAGCTGTGCAACCGGCGTCTGTGC 1046
Db 721 GCGCTCACCACTAGTACGCGCGCGCTTCTGGGTACAGCTGTGCAACCGGCGTCTGTGC 780
QY 1047 CTCTTCTCTGGAAGGGGCGTGTGAGTCTTCAGTATGTTCCGCGCTCTTCGCAAC 1106
Db 781 CTCTTCTCTGGAAGGGGCGTGTGAGTCTTCAGTATGTTCCGCGCTCTTCGCAAC 840
QY 1107 CTCTTGAACCAAAAGCGCAGAGCTGACGACAGAGAGAGGGGGCTCACTCTTATCTTC 1166
Db 841 CTCTTGAACCAAAAGCGCAGAGCTGACGACAGAGAGAGGGGGCTCACTCTTATCTTC 900
QY 1167 GCGGACCACTGACAGAGAGCGCTCTCCAGATTAAATGATTCACCTAACCTG 1226
Db 901 GCGGACCACTGACAGAGAGCGCTCTCCAGATTAAATGATTCACCTAACCTG 960
QY 1227 TGA 1229
Db 961 TGA 963

```

```

RESULT 2
LOCUS DQ050096 769 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes HCl8097 gene, VIRUAL TRANSCRIPT, partial sequence,

```

```

ACCESSION DQ050096
VERSION DQ050096.1 GI:66903295
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
MAMMALIA; EUTHERIA; EUARCHONTOMORPHIA; PRIMATES; CATARRHINI; HOMINIDAE; PAN.
REFERENCE 1 (bases 1 to 769)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B.,
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D.,
White,T.J., Sniinsky,J.J., Adams,M.D. and Cargill,M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
(ser) PLoS Biol. 3 (6), E170 (2005)
JOURNAL 15869325
PUBMED 2 (bases 1 to 769)
REFERENCE Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B.,
AUTHORS Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D.,
White,T.J., Sniinsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source Location/Qualifiers
gene
ORIGIN
Query Match 49.4%; Score 701.6; DB 11; Length 769;
Best Local Similarity 91.5%; Pred. No. 7.9e-169;
Matches 704; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 267 ATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCCGGAGTCCGAGGCTTC 326
Db 1 ATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCCGGAGTCCGAGGCTTC 60
QY 327 AGCGTTCACTGTGCTACATCGTTATTTAGTGTGTTTGGCTCTAGACAGAGCTTCGTCTC 386
Db 61 AGCGTTCACTGTGCTACATCGTTATTTAGTGTGTTTGGCTCTAGACAGAGCTTCGTCTC 120
QY 387 ATCTTGGCCGGGGATCCGTGGCCCATCGCGCTGGTGTGTTGGTGTGAGTCTTCTCACT 446
Db 121 ATCTTGGCCGGGGATCCGTGGCCCATCGCGCTGGTGTGTTGGTGTGAGTCTTCTCACT 180
QY 447 CTGTTATAGGCGCAGAAATTTGGCTGTGCACTTCACTGACAGAAATGTTTCGTGGTACA 506
Db 181 CTGTTATAGGCGCAGAAATTTGGCTGTGCACTTCACTGACAGAAATGTTTCGTGGTACA 240
QY 507 GTGAACACCAACATCTTCAAAAGCCTTCAAGCGCAGCGCGCTTACAGCCGTCGTGGT 566
Db 241 GTGAACACCAACATCTTCAAAAGCCTTCAAGCGCAGCGCGCTTACAGCCGTCGTGGT 300
QY 567 CTGCTGTGGGCGCTGAGGGCATTATATTTATACCTGACAGGGACCCCACTGATCAGCTG 626
Db 301 CTGCTGTGGGCGCTGAGGGCATTATATTTATACCTGACAGGGACCCCACTGATCAGCTG 360
QY 627 AACGAGACCATTTGACTACAAAGAGAGTTCACCTGGCGTCTGAAGAAGATTACGCGCG 686
Db 361 AACGAGACCATTTGACTACAAAGAGAGTTCACCTGGCGTCTGAAGAAGATTACGCGCG 420
QY 687 GAGTACGGAACGCACTGAGAAAGGGGCTGCGGACCCAGTGTCTTACCTGGCGGAGAA 746
Db 421 GAGTACGGAACGCACTGAGAAAGGGGCTGCGGACCCAGTGTCTTACCTGGCGGAGAA 480
QY 747 TTCAACACGAGTAGGCTTGGCGGCTGTACCAACAGTACCACTGGCGGAGACCTACAGCC 806

```

[illegible]

TITLE
JOURNAL

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGGAGAGATCCAGACCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 10.0 and subcloning to
Rot = 105.2. Second strand cDNA was prepared with the primer
adapter of sequence [5'
GAGAGAGATCTTCAGTTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda PLoC I. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B.

FEATURES
source
Location/Qualifiers
1. 1594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:9030623N16"
/db_xref="taxon:10090"
/clone="9030623N16"
/sex="male"
/tissue_type="colon"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
144..1106
/note="unnamed protein product; hypothetical protein
(evidence: ProCrest,decoder,NCBI CDS
Predictor,Longest-ORF)
putative"
/codon_start=1
/protein_id="BAB11281.1"
/db_xref="GI:12858339"
/translation="MTAMDVLEPPYPPRPAAGSVPLLVILVPLSLAASPLTIRE
IRKHSRPMVLVRVLISLFGAEIIVAVPSDGVFGRWNTSYAFSPSRVQVVGLE
VGLAGVNIILTRGPRQGLNITDIENRFTPLRNDYRLEVHALKEKLPDVLIAE
FTPSPCGLHGYHVALAAATLWVAFCEWIIINALISMPALPYGIGALLTGAFTI
RGVPAFASISVPLCHRFISAVITTPYVAGSPMLTANGLISLGGAVILHTRP
ALNSFDLSVYDCSNQAKNSPLTIANPQHQLKSPDLNTTTL"

ORIGIN
Query Match 48.8%; Score 693; DB 4; Length 1594;
Best Local Similarity 75.9%; Pred. No. 1,46-166;
Matches 914; Conservative 0; Mismatches 275; Indels 16; Gaps 4;

DB 230 CTGCGCGCTGCTGTCAGACTGGCGCGGCGCTGTCAGATGACCTGTGAAACGGCGTACT 289
107 CTCGCTCGCTCGGAACAGTTCTCCAACTTCTTGACAGATGACTCGTGGACGGCGTCT 166
290 GCCTTTTACCGCAGCGCCGGGAGTGGCGGAGGCTTACGCTTCACTGCTCATCGTTAT 349
DB 167 ACCCTTTTACCGCAGCGCGGAGTGGCGGAGGCTTACGCTTCACTGCTCATCGTTAT 226

QY 350 TCTAGTGTGTTGGCTCTAGCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCA 409
 DB 227 CTTGGTGTCTTGTCTTGGCTGTAGCTTCTGCTCATCTTGGCATCCGTGGGGA 286
 QY 410 CTCGGCTGTGTTGGTGTGAGAGTTCTTCTCACTGTCTTCACTAGCGCAGAAATTGT 469
 DB 287 CTCGGCTGTGTTGGTGTGAGAGTTCTTCTCACTGTCTTCACTAGCGCAGAAATTGT 346
 QY 470 GGCTGTGCACTTCACTGAGCAAGATGTTGTGTGCTACAGTGAACCAACAATCTTACAA 529
 DB 347 GGCTGTGCACTTCACTGAGCAAGATGTTGTGTGCTACAGTGAACCAACAATCTTACAA 406
 QY 530 AGCTTCAAGCGAGCGCGGTTTACACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 589
 DB 407 AGCTTCAAGCGAGCGCGGTTTACACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 466
 QY 590 TATATTTAATCACTACAGCAAGCACTGATGATGATGATGATGATGATGATGATGATGAT 649
 DB 467 TATATTTAATCACTACAGCAAGCACTGATGATGATGATGATGATGATGATGATGATGAT 526
 QY 650 GCACTTCACTGCTGTGTAAGAGATTAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 709
 DB 527 GCTTTCACCTGCTGTGTAAGAGATTAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 586
 QY 710 GGGGCTGCGCGGAGCCAGTGTCTAAGCTGTGCGAGAGTTCAACAGAGTACCTTGGCGG 769
 DB 587 GGGGCTGCGCGGAGCCAGTGTCTAAGCTGTGCGAGAGTTCAACAGAGTACCTTGGCGG 646
 QY 770 CCTGTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
 DB 647 GCTGTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 706
 QY 830 CTGCTTGTGCTGTCTCTCAAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 889
 DB 707 CTGCTTGTGCTGTCTCTCAAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 766
 QY 890 GGCACGTGAGCAACCGGAGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 949
 DB 767 GGCACGTGAGCAACCGGAGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 826
 QY 950 TACGTCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1009
 DB 827 TACGTCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 886
 QY 1010 CGCTTGTGCTGTCTCTCAAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1069
 DB 887 CGCTTGTGCTGTCTCTCAAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 946
 QY 1070 GAGTCTCAGATGTTGGGCGGAGCTTGTGCAACCTTGTGCAACCAAGCCCAAGCA 1129
 DB 947 GATTTCTCACTAATCTGCGCCAGCGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1006
 QY 1130 CT---GAGCCAGAGAGAGGAGGAGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1186
 DB 1007 CTGTAGCAACCAAGCTGTGAGAACTCACTGTGCAACCTGTGCAACCAAGCCCAAGCA 1066
 QY 1187 GGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1246
 DB 1067 GTTGAAGAGTCCAGACTTAA---TATTTACCACTCTCTGTGAAATTAATCTGACTCCGGA 1123
 QY 1247 CTCTTCTGCGGCTGTGAGATCTGCAAGCGGAGAGAGTGTGCGGCGGAGCTGTG-GGCCA 1305
 DB 1124 TTTCTACCTCTCTTGTGAGACCCCTAATCACTGTGAGAGTGTGTTAAGAGCGGTGTGCAAGA 1183
 QY 1306 GAGAGAGTCCAGAGAGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1365
 DB 1184 GCAATAGCTGTGTGAGAACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1243
 QY 1366 CC-----AGGAGAGTCTCTGTGAGAGATCTGTGAAATTAATCTTGTCTTGT 1416
 DB 1244 CTTTGTCTTGAAGAAATGATTTCTGAGAGAACTGTGAAATTAATCTTGTCTTGT 1303

QY 1417 TTTT 1421
 DB 1304 TTTT 1308
 RESULT 4
 AK008816
 LOCUS
 DEFINITION
 AK008816 1309 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male stomach cDNA, RIKEN full-length enriched
 library, clone:2210403N03 product:hypothetical protein, full insert
 sequence.
 ACCESSION
 AK008816
 VERSION
 AK008816.1 GI:12843238
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 1042159
 THE RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 THE PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1309)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hirooka, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, H.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishii, K.,
 Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Teijima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
 Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submision
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.jp),

REFERENCE 1 (bases 1 to 619)
 AUTHORS Stevens, M., Wei, C., Gross, S.S., McPherson, J. and Brent, M.R.
 TITLE Exhaustive RT-PCR and sequencing of all novel TWINSKAN predictions
 JOURNAL Unpublished (2005)
 COMMENT Contact: Brent, MR
 Laboratory of Computational Genomics
 Washington University
 Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
 Tel: (314) 935-6621
 Email: brent@ese.wustl.edu

All products are PCR'd from first-strand cDNA made from pooled mRNA sources. PCR product is sequenced directly, typically from the PCR primers. All submitted sequences produced a high quality, spliced alignment to the genome or assembled into a contig that aligned to the genome. This alignment was not necessarily to the targeted locus. Legitimate, spliced sequences that resulted from mispriming to non-targeted cDNAs were also submitted. All data on all experiments can be found at <http://genes.cse.wustl.edu/rtdb/>.

FEATURES
 source Location/Qualifiers

1..619
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGCL.5.1.1.1.A11"
 /issue_type="pooled"
 /clone_lib="NIH_MGC_331"
 /note="Tissues: adrenal gland, bone marrow, brain cerebellum, brain (whole), fetal brain, fetal liver, heart, kidney, liver, lung (whole), placenta prostate, salivary gland, skeletal muscle, testis, thymus, thyroid gland trachea, uterus, spinal cord. First-stranded cDNA made from pooled mRNA sources. PCR product is sequencing directly, typically from the PCR primers."

ORIGIN

Query Match 43.3%; Score 615.8; DB 8; Length 619;
 Best Local Similarity 99.7%; Pred. No. 8.2e-147;
 Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

302 CCAAGCCCGGCGATGCCGAGGCTTCAAGCGTTCATGCTCATGTTATTGTTGTTT 361
 DB 619 CCAAGCCCGGCGATGCCGAGGCTTCAAGCGTTCATGCTCATGTTATTGTTGTTT 560
 QY 362 GGCTCTAGAGAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGGGCGACCTCGCGCTGTT 421
 DB 559 GGCTCTAGAGAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGGGCGACCTCGCGCTGTT 500
 QY 422 TTGGTTGTAGAGTTCTTCTCAGTCTGTTCAATAGCGGAGAAATTGTGCTGTGCACTT 481
 DB 499 TTGGTTGTAGAGTTCTTCTCAGTCTGTTCAATAGCGGAGAAATTGTGCTGTGCACTT 440
 QY 482 CAGTGCAGATGTTGTGTGGTACAGTGAACACCAACAATCTTACAAAGCTTCAAGCCG 541
 DB 439 CAGTGCAGATGTTGTGTGGTACAGTGAACACCAACAATCTTACAAAGCTTCAAGCCG 380
 QY 542 AAGCGCGTTTACAGCCGCTGTGCTGCTCGGCGGCTCGAGGCGGATTAATTAATTA 601
 DB 379 AAGCGCGTTTACAGCCGCTGTGCTGCTCGGCGGCTCGAGGCGGATTAATTAATTA 320
 QY 602 CACAGGAGCCCGAGTGCATCAGTGAACGAGACCATTTGACACAGAGAGTTCACTG 661
 DB 319 CACAGGAGCCCGAGTGCATCAGTGAACGAGACCATTTGACACAGAGAGTTCACTG 260
 QY 662 GCGTCTGAAGAGATTAAGCGCGGAGTACGCGAAAGCACTGAGAGAGGCGCTGCGGA 721
 DB 259 GCGTCTGAAGAGATTAAGCGCGGAGTACGCGAAAGCACTGAGAGAGGCGCTGCGGA 200
 QY 722 CCCAGGCTCTACCTGCGGAGAGAGTTCAACGAGATAGCCCTTGGGCGCTGTACACACA 781
 DB 199 CCCAGGCTCTACCTGCGGAGAGAGTTCAACGAGATAGCCCTTGGGCGCTGTACACACA 140
 QY 782 GTACCACTTGCGGAGACACTACGCTCTGCGCACGCTATGGGTGGCGTTGCTTGCGCT 841

DB 139 GTACCACTTGCGGAGACACTACGCTCTGCGCACGCTATGGGTGGCGTTGCTTGCGCT 80
 QY 842 CCTTCCAAAGCTGCTCTTCCAGCCCGGCGCTTACGAGAGGCTTGGCACTGTGAC 901
 DB 79 CCTTCCAAAGCTGCTCTTCCAGCCCGGCGCTTACGAGAGGCTTGGCACTGTGAC 20
 QY 902 CACCGAGGCTTGGCGCTC 920
 DB 19 CACCGAGGCTTGGCGCTC 1

RESULT 6
 DN831247
 LOCUS
 DEFINITION MGCL.5.1.1.1.A11.F.1 NIH_MGC_331 Homo sapiens cDNA clone
 ACCESSION MGCL.5.1.1.1.A11, mRNA sequence.
 VERSION DN831247
 KEYWORDS DN831247.1 GI:62639816
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE 1 (bases 1 to 617)
 AUTHORS Stevens, M., Wei, C., Gross, S.S., McPherson, J. and Brent, M.R.
 TITLE Exhaustive RT-PCR and sequencing of all novel TWINSKAN predictions
 JOURNAL Unpublished (2005)
 COMMENT Contact: Brent, MR
 Laboratory of Computational Genomics
 Washington University
 Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
 Tel: (314) 935-6621

Email: brent@ese.wustl.edu
 All products are PCR'd from first-strand cDNA made from pooled mRNA sources. PCR product is sequenced directly, typically from the PCR primers. All submitted sequences produced a high quality, spliced alignment to the genome or assembled into a contig that aligned to the genome. This alignment was not necessarily to the targeted locus. Legitimate, spliced sequences that resulted from mispriming to non-targeted cDNAs were also submitted. All data on all experiments can be found at <http://genes.cse.wustl.edu/rtdb/>.

FEATURES
 source Location/Qualifiers

1..617
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGCL.5.1.1.1.A11"
 /issue_type="pooled"
 /clone_lib="NIH_MGC_331"
 /note="Tissues: adrenal gland, bone marrow, brain cerebellum, brain (whole), fetal brain, fetal liver, heart, kidney, liver, lung (whole), placenta prostate, salivary gland, skeletal muscle, testis, thymus, thyroid gland trachea, uterus, spinal cord. First-stranded cDNA made from pooled mRNA sources. PCR product is sequencing directly, typically from the PCR primers."

ORIGIN

Query Match 43.0%; Score 610.6; DB 8; Length 617;
 Best Local Similarity 99.4%; Pred. No. 1.8e-145;
 Matches 613; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

341 CATCGTTATCTAGTGTGTTTGGCTTACAGCAAGCTTCTGCTCATCTTGGCGGAGT 400
 DB 1 CATCGTTATCTAGTGTGTTTGGCTTACAGCAAGCTTCTGCTCATCTTGGCGGAGT 60
 QY 401 CCGTGGCACTGCGCGTGTGTTTGGTGTGAGAGTTCTTCAAGCTTCAATAGGCGC 460
 DB 61 CCGTGGCACTGCGCGTGTGTTTGGTGTGAGAGTTCTTCAAGCTTCAATAGGCGC 120

QY 461 AGAATTTGGGCTGTGACCTTCACTGACAGAAATGTTCTGTGGTACAGTGAACCAAC 520
 DB 121 AGAATTTGGGCTGTGACCTTCACTGACAGAAATGTTCTGTGGTACAGTGAACCAAC 180
 QY 521 ATCTTCAAAAGCCTTACAGCAGCGCGCTTACAGCCCGTGTGCTGTGCTG 580
 DB 181 ATCTTCAAAAGCCTTACAGCAGCGCGCTTACAGCCCGTGTGCTGTGCTG 240
 QY 581 GAGGGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 640
 DB 241 GAGGGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
 QY 641 CTACAAAGAGCACTTACCTGTGGCTGAAAGAAATTAAGCCCGGAGTACGGAACG 700
 DB 301 CTACAAAGAGCACTTACCTGTGGCTGAAAGAAATTAAGCCCGGAGTACGGAACG 360
 QY 701 ACTGAGAAAGGCGCTGCGAGCCAGTGTCTTACCTGTGGAGAAAGTTCAACCGAGTAG 760
 DB 361 ACTGAGAAAGGCGCTGCGAGCCAGTGTCTTACCTGTGGAGAAAGTTCAACCGAGTAG 420
 QY 761 CCTTGGCGCTTACCAACCAAGTACCACTGTGGCGGACACTAGCCTGGCCAGCTATG 820
 DB 421 CCTTGGCGCTTACCAACCAAGTACCACTGTGGCGGACACTAGCCTGGCCAGCTATG 480
 QY 821 GGTGGCTTGTCTTGTGCTTCTCTCTCAACGTGTGCTCTCAACGCGCGCGCTCTA 880
 DB 481 GGTGGCTTGTCTTGTGCTTCTCTCTCAACGTGTGCTCTCAACGCGCGCGCTCTA 540
 QY 881 CGAGAGCCTGACCTGTGACCAACCGAGCCTTGGCCTCTGCGGGCTTTCGCTTGGC 940
 DB 541 CGAGAGCCTGACCTGTGACCAACCGAGCCTTGGCCTCTGCGGGCTTTCGCTTGGC 600
 QY 941 CTCATCTCTAGCGTGC 957
 DB 601 CTCATCTCTAGCGTGC 617

RESULT 7
 DN930503/c 602 bp mRNA linear EST 27-APR-2005
 LOCUS
 DEFINITION MGCL.5.1.2.1.A11.R.1 NIH MGCL 331 Homo sapiens cDNA clone
 ACCESSION DN930503
 VERSION DN930503.1 GI:62924189
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Euteleostomi; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 602)
 Stevens, M., Wei, C., Gross, S.S., McPherson, J. and Brent, M.R.
 Exhaustive RT-PCR and sequencing of all novel TWINSKAN predictions
 in human
 Unpublished (2005)
 CONTACT: Brent, M.
 Laboratory of Computational Genomics
 Washington University
 Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
 Tel: (314) 935-6621

FEATURES
 Source
 1..602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 Email: brent@case.wustl.edu
 All products are PCR'd from first-strand cDNA made from pooled mRNA
 sources. PCR product is sequenced directly, typically from the PCR
 primers. All submitted sequences produced a high quality, spliced
 alignment to the genome or assembled into a contig that aligned to
 the genome. This alignment was not necessarily to the targeted
 locus. Legitimate, spliced sequences that resulted from mispriming
 to non-targeted cDNAs were also submitted. All data on all
 experiments can be found at <http://genes.case.wustl.edu/tld/>.
 Location/Qualifiers

/db_xref="taxon:9606"
 /clone="MGCL.5.1.2.1.A11"
 /issue_type="pooled"
 /clone_id="NIH_MGCL_331"
 /note="tissues: adrenal gland, bone marrow, brain
 cerebellum, brain (whole), fetal brain, fetal liver,
 heart, kidney, liver, lung (whole), placenta, prostate,
 salivary gland, skeletal muscle, testis, thymus, thyroid
 gland, trachea, uterus, spinal cord. First-strand cDNA
 made from pooled mRNA sources. PCR product is sequencing
 directly, typically from the PCR primers."

ORIGIN
 Query Match 42.4%; Score 602; DB 8; Length 602;
 Best Local Similarity 100.0%; Pred. No. 2.9e-143;
 Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 CCAGCCCGGATGCGGAGGCTTACAGCCTTCACTGCTCATGTTATTTAGTGT 361
 DB 602 CCAGCCCGGATGCGGAGGCTTACAGCCTTCACTGCTCATGTTATTTAGTGT 543
 QY 362 GGTCTTACAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGGCCACTGGGCTG 421
 DB 542 GGTCTTACAGCAAGCTTCTGCTCATCTTGGCGGAGATCCGTGGCCACTGGGCTG 483
 QY 422 TTGGTGTGAGAGTCTTCTCACTGCTCATGAGGCGAGAAATGTGGCTGCACTT 481
 DB 482 TTGGTGTGAGAGTCTTCTCACTGCTCATGAGGCGAGAAATGTGGCTGCACTT 423
 QY 482 CAGTGAAGATGTTGTGTGTGATCAGTGAACCAACATCTCTACAAAGCCTTACGCGC 541
 DB 422 CAGTGAAGATGTTGTGTGTGATCAGTGAACCAACATCTCTACAAAGCCTTACGCGC 363
 QY 542 AGCGCGGTTACAGCCCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 601
 DB 362 AGCGCGGTTACAGCCCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 303
 QY 602 CACAGGAGACCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 661
 DB 302 CACAGGAGACCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 243
 QY 662 GCGTCTGAAGAGAAATTAACCGCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 721
 DB 242 GCGTCTGAAGAGAAATTAACCGCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 183
 QY 722 CCAAGTCTTACTGTGCGGAGAAATTAACCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 781
 DB 182 CCAAGTCTTACTGTGCGGAGAAATTAACCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 123
 QY 782 GTACCACTGTGCGGAGAAATTAACCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 841
 DB 122 GTACCACTGTGCGGAGAAATTAACCGGAGTACCGGAAAGCACTGAGAAAGGCGCTGCGGA 63
 QY 842 CTTCTCAAGTGTGCTTCTCAACGCGCGCGCTTCTTACGAGAGGCTTGGCATGCTGAC 901
 DB 62 CTTCTCAAGTGTGCTTCTCAACGCGCGCGCTTCTTACGAGAGGCTTGGCATGCTGAC 3
 QY 902 CA 903
 DB 2 CA 1

RESULT 8
 DN930502 622 bp mRNA linear EST 27-APR-2005
 LOCUS
 DEFINITION MGCL.5.1.2.1.A11.F.1 NIH MGCL 331 Homo sapiens cDNA clone
 ACCESSION DN930502
 VERSION DN930502.1 GI:62924188
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo;
 1 (bases 1 to 622)
 Stevens, M., Wei, C., Gross, S.S., McPherson, J. and Brent, M.R.
 Exhaustive RT-PCR and sequencing of all novel TWINSKAN predictions
 in human
 Unpublished (2005)
 Contact: Brent, MR
 Laboratory of Computational Genomics
 Washington University
 Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
 Tel: (314) 935-6621
 Email: bren@cse.wustl.edu
 All products are PCR'd from first-strand cDNA made from pooled mRNA
 sources. PCR product is sequenced directly, typically from the PCR
 primers. All submitted sequences produced a high quality, spliced
 alignment to the genome or assembled into a contig that aligned to
 the genome. This alignment was not necessarily to the targeted
 locus. Legitimate, spliced sequences that resulted from mispriming
 to non-targeted cDNAs were also submitted. All data on all
 experiments can be found at <http://genes.cse.wustl.edu/rtdb/>.
 location/qualifiers
 1..622
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC1.5.1.2.1.All"
 /tissue_type="pooled"
 /clone_id="NIH_MGC_331"
 /note="Tissues: adrenal gland, bone marrow, brain
 cerebellum, brain (whole), fetal brain, fetal liver,
 heart, kidney, liver, lung (whole), placenta proctate,
 salivary gland, skeletal muscle, testis, thymus, thyroid
 gland trachea, uterus, spinal cord. First-strand cDNA
 made from pooled mRNA sources. PCR product is sequencing
 directly, typically from the PCR primers."

ORIGIN

Query Match 42.3%; Score 600.8; DB 8; Length 622;
 Best Local Similarity 98.1%; Pred. No. 5,8e-143;
 Matches 608; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

332 TCACATGCTCATGTTATTTCTAGTGTGTTGGCTCTAGACGACAGCTTCTGCTCATCTT 391
 2 TTACATGCTCATGTTATTTCTAGTGTGTTGGCTCTAGACGACAGCTTCTGCTCATCTT 61

392 GCCGAGGATCCGTCGACCTGCGCTGTTGGTGTGAGAGCTCTCTGACGCTGTT 451
 62 TAAGGCGATCCGTCGACCTGCGCTGTTGGTGTGAGAGCTCTCTGACGCTGTT 121

452 CATAGCGCGAAGATTTGCTGTGACCTTCACTGACAGATGTTGCTGAGTACAGTGA 511
 122 CATAGCGCGAAGATTTGCTGTGACCTTCACTGACAGATGTTGCTGAGTACAGTGA 181

512 CACCAACATCTCTAAGAGCTTCAAGCGGACGCGGTTACAGCCCGTGTGCTGCT 571
 182 CACCAACATCTCTAAGAGCTTCAAGCGGACGCGGTTACAGCCCGTGTGCTGCT 241

572 CCGTGGCTGAGAGGATTAATTTACCTCAACAGGAGCCCGATGATCACTGAACA 631
 242 CCGTGGCTGAGAGGATTAATTTACCTCAACAGGAGCCCGATGATCACTGAACA 301

632 GACCATTTGATCAACAGAGATTTCACTGCGGCTGTGAAGAAGATTAAGCGCGAGTA 691
 302 GACCATTTGATCAACAGAGATTTCACTGCGGCTGTGAAGAAGATTAAGCGCGAGTA 361

692 CCGGAACGACTGAGAGAGGCTGCGGACCGAGTCTTACTGCGGAGAGAGTTAC 751
 362 CCGGAACGACTGAGAGAGGCTGCGGACCGAGTCTTACTGCGGAGAGAGTTAC 421

752 ACCGAGTACCCCTTGGGCTGTACCAACAGTACCACTGCGGAGCACTACGCTCGAC 811
 422 ACCGAGTACCCCTTGGGCTGTACCAACAGTACCACTGCGGAGCACTACGCTCGAC 481

ORIGIN

Query Match 37.3%; Score 529.6; DB 1; Length 537;
 Best Local Similarity 99.3%; Pred. No. 1.1e-124;
 Matches 532; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

812 CACGCTAGGAGTGGGCTTGTGCTCTCTCTCCACAGTGTCTCTCCACGCGGAC 871
 482 CACGCTAGGAGTGGGCTTGTGCTCTCTCTCCACAGTGTCTCTCCACGCGGAC 541

872 CCGCTCTACAGAGGCTTGGGCTGTGACCAACCGAGCTTGGGCTTGGGCTT 931
 542 CCGCTCTACAGAGGCTTGGGCTGTGACCAACCGAGCTTGGGCTTGGGCTT 601

932 CCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 951
 602 CCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 621

RESULT 9
 A1821606/c
 LOCUS
 DEFINITION
 A1821606 537 bp mRNA linear EST 13-DEC-1999
 nk08a11.x5 NCI CGAP C02 Homo sapiens cDNA clone IMAGE:1012892 3'
 similar to contig1ns.TARI.t3 TARI MER22 repetitive element ; mRNA
 sequence.
 A1821606
 A1821606.1 GI:5440685
 EST.
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 537)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: nk08a11.y5
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmerit-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc., David B. Kitzman,
 Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILMUT at:
www.bio.lnl.gov/bbtrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
 Original clone citation: National Cancer Institute, Cancer Genome
 Anatomy Project (CGAP), Tumor Gene Index
 This read has been verified (found to hit its original self in the
 correct orientation)
 Insert Length: 741 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 459.
 Location/Qualifiers
 1..537
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1012892"
 /tissue_type="tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_id="NCI CGAP C02"
 /note="Organ: colon; Vector: Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence:
 5' GAATTCGACGAG 3' 3' adaptor sequence: 5'
 CTCGAGTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

QY 886 GCGTGGCACTGCTGACCAACCGAGCCTTGGCGCTTTCGGGGGTTCCTGCGCTTGGCCCTCCA 945
 DB 537 GCGTGGCACTGCTGACCAACCGAGCCTTGGCGCTTTCGGGGGTTCCTGCGCTTGGCCCTCCA 478
 QY 946 TCTCTAGCGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 1005
 DB 477 TCTCTAGCGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 418
 QY 1006 GCGCGCGCTTCTGGGTGACGCTGGGCAACCGGCGCTTGGCGCTTCTTCTTCTGGAGGGGCGG 1065
 DB 417 GCGCGCGCTTCTGGGTGACGCTGGGCAACCGGCGCTTGGCGCTTCTTCTTCTGGAGGGGCGG 358
 QY 1066 TGGTGAATCTCCAGTATGTTTCGGCGCCGAGCGCTCTTGGCAACCTTGGAGCAAAAGCGCA 1125
 DB 357 TGGTGAATCTCCAGTATGTTTCGGCGCCGAGCGCTCTTGGCAACCTTGGAGCAAAAGCGCA 298
 QY 1126 AGGACTGACAGCAG 1185
 DB 297 AGGACTGACAGCAG 238
 QY 1186 AGGCGCGCTTCCAGACTTAATGATATCACTAACTTGTAGAGGGGAGCCCAATCTGG 1245
 DB 237 AGGCGCGCTTCCAGACTTAATGATATCACTAACTTGTAGAGGGGAGCCCAATCTGG 178
 QY 1246 ACTCTTCTCCCGCTTGGAGCATGCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1305
 DB 177 ACTCTTCTCCCGCTTGGAGCATGCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 118
 QY 1306 GAGAGCTCCAG 1365
 DB 117 GAGAGCTCCAG 58
 QY 1366 CAGAGGAAGTCTCTGGGGGAGATCTGTAAATAAACCCTTTTCTTTGTTTTT 1421
 DB 57 CAGAGGAAGTCTCTGGGGGAGATCTGTAAATAAACCCTTTTCTTTGTTTTT 2

RESULT 10
 CK971974 717 bp mRNA linear EST 16-MAR-2004
 LOCUS 4087846 BARC 9BOV Bos taurus cDNA clone 9BOV4_E23 5', mRNA
 DEFINITION
 ACCESSION
 VERSION CK971974
 KEYWORDS
 SOURCE EST.
 ORGANISM Bos taurus (cow)
 Bos taurus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 717)
 Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
 G.P., Bosak, S., Rubenfield, M., and Gabyarre, L.C.
 Production of EST from cDNA libraries derived from immunologically
 activated bovine gut
 Unpublished (2004)
 Contact: Tad S. Sonstegard
 Bovine Functional Genomics Laboratory
 Animal and Natural Resources Institute
 Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048416
 Fax: 3015048416
 Email: tads@anr.ars.usda.gov

JOURNAL
 COMMENT
 FEATURES
 source
 1. 717
 Location/Qualifiers
 High quality sequence stop: 717.
 Seg primer: CCGAGTCAAGAGCTGTAAAGC
 Plate: 4 row: B column: 23
 by cross_match using options -mismatch 12 -minscore 18
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim alt - -trim fasta. Vector identified
 by cross_match using options -mismatch 12 -minscore 18

/mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="9BOV4_E23"
 /sex="Male"
 /tissue_type="Pooled"
 /dev_stage="Multiple"
 /lab_host="DH10B T1 phage resistant"
 /clone_1lb="BARC 9BOV"
 /notes="Organ: Abomasum; Vector: pagen-1; Site: 1; ScoreV;
 Site: 2; Note: Equimolar amounts of mRNA extracted from
 fundic and pyloric abomasums of 18 and 21 week old steers.
 Exposure to Osteragia osteragi was initiated at 15 weeks
 of age. fundic and pyloric abomasum"

ORIGIN
 Query Match 37.0%; Score 526.4; DB 7; Length 717;
 Best local similarity 83.7%; Pred. No. 7.5e-124;
 Matches 596; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 250 CTGGCGCGGTCGACAGTACACCTGTGGAAGGGCGTACTGCTTTTACCCAGGCC 309
 DB 6 CTGAGCGCGGTGACAGTACCTGTGAGATGTGTGCGCTTCTACCTCAGCC 65
 QY 310 GGCATGCGCGAGCTTCAGCGTTCACCTGCTCATGTTATCTAGTGTGCTCTAG 369
 DB 66 GGCATGCGCGCGCTTCAGTGTGCTGCTCATGTTATCTAGTGTGCTCTAG 125
 QY 370 CAGCAAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCACTGCGCTGTTGTTGG 429
 DB 126 CCGCAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCACTGCGCTGTTGTTGG 185
 QY 430 TGAAGTCTTCTCAAGTCTGTTATGAGGCGAGAAATGAGGCTGAGCACTTCA 489
 DB 186 TGAAGTCTTCTCAAGTCTGTTATGAGGCGAGAAATGAGGCGCACTTCA 245
 QY 490 AATGTTCTGTGTTACAGTGAACCAACCAATCTTCAAAAGCTTCAAGCGCGCG 549
 DB 246 AATGTTCTGTGTTACAGTGAACCAACCAATCTTCAAAAGCTTCAAGCGCG 305
 QY 550 TTACAGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
 DB 306 TCCGAGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
 QY 610 CCCGATGCTACAGTGAACCAACCAATCTTCAAAAGCTTCAAGCGCGCG 669
 DB 366 ACCGATGCTACAGTGAACCAACCAATCTTCAAAAGCTTCAAGCGCGCG 425
 QY 670 AAGGATTTACGCGCGGAGTGAACCAACCAATCTTCAAAAGCTTCAAGCGCGCG 729
 DB 426 GCCAAATCTATGCAAGGCGCTGACGAGGCGCTGAGAGGCGCTGCCAACCGGTTTC 485
 QY 730 TCTACTGCGGAGAGAGTTCACACCGAGTACGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
 DB 486 TCTATCTGGGAGAGAGTTCACACCGAGTACGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCT 545
 QY 790 TGGCGGAGCACTACGCTTGGCGCAACGCTATAGGAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 849
 DB 546 TGGCGGAGCACTACGCTTGGCGCAACGCTATAGGAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 605
 QY 850 ACGTGTGCTTCTACAGCGCGCGCTTGGCGCAACGCTATAGGAGGCTTGTGCTTGTGCTTGTGCTTGTGCT 909
 DB 606 ACATGTGCTTCTACAGCGCGCGCTTGGCGCAACGCTATAGGAGGCTTGTGCTTGTGCTTGTGCTTGTGCT 665
 QY 910 CTTTGGGCTTCTTGGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 961
 DB 666 CTTTGGGCTTCTTGGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 717

RESULT 11
 CK979918 703 bp mRNA linear EST 16-MAR-2004
 LOCUS 411631 BARC 9BOV Bos taurus cDNA clone 9BOV40_D22 5', mRNA
 DEFINITION

sequence.
 ACCESSION CK979918 GI:45497898
 VERSION CK979918.1
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 703)
 Sonnegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Hartney, G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
 Production of EST from cDNA libraries derived from immunologically activated bovine gut
 Unpublished (2004)
 JOURNAL Contact: Tad S. Sonnegard
 Bovine Functional Genomics Laboratory
 Animal and Natural Resources Institute
 Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048416
 Fax: 3015048414
 Email: tads@anri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim_alt "-trim_fasta. Vector identified by cross_match using options -mismatch 12 -minscore 18
 Plate: 40 row: D column: 22
 Seq primer: CCCAGTCACGACGTTGAACG
 High quality sequence stop: 703.
 Location/Qualifiers
 1..703
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="9BOV40_D22"
 /sex="Male"
 /tissue_type="Pooled"
 /dev_stage="Multiple"
 /lab_host="DH10B T1 phage resistant"
 /clone_1lb="BARC 9BOV"
 /note="Organ: Abomasum; Vector: Pagen-1; Site 1: EcorV; Site 2: NotI; Equimolar amounts of mRNA extracted from fundic and pyloric abomasum of 18 and 21 week old steers. Exposure to Osteragia osteraggi was initiated at 15 weeks of age. fundic and pyloric abomasum"
 ORIGIN
 Query Match 36.7%; Score 522.2; DB 7; Length 703;
 Best Local Similarity 83.9%; Pred. No. 9e-123;
 Matches 590; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
 QY 288 CTGCGCTTTTATCCCGCAGCCCGGAGTCCGAGGCTTCAAGCGTTCACTGCTCATCGTT 347
 DB 1 CTGCGCTTTTATCCCGCAGCCCGGAGTCCGAGGCTTCAAGCGTTCACTGCTCATCGTT 60
 QY 348 ATTCTAGTGTGTTTGGCTTCAAGAGAGAGTCCGCTCATCTTCCGCGGAGTCCGTTGC 407
 DB 61 ATTCTAGTGTGTTTGGCTTCAAGAGAGAGTCCGCTCATCTTCCGCGGAGTCCGTTGC 120
 QY 408 CACTCGCGCTGTTTGGTGTGAGAGTCTTCTCACTGCTGTTCAATAGGCGCAGAAATT 467
 DB 121 CACTCGCGAGTGTGTTGGTGTGAGAGTCTTCTCACTGCTGTTCAATAGGAGCAAAATT 180
 QY 468 GTGCGTGTGCACTTCAAGTCAAGATGTTCTGGGTACAGTGAACAACAACATCTTAC 527
 DB 181 GTGCGCGGCACTTCAAGTCAAGATGTTCTGGGTACAGTGAACAACAACATCTTAC 240
 QY 528 AAAGCCTTACGCGAGCGCGCTTACAGCCCGGTGCGTGTGCGTGTGCGCTTGAAGGCG 587
 DB 241 AAAGCCTTACGCGAGCGCGCTTCAAGCGCGCTTCAAGCGCGCTTGAAGGCGCTTGAAGG 300
 QY 588 ATTATATTTACCTACAGAGGACCCCAAGTGCATCAGTGAACGAGACCATTTGACTTAAAC 647

DB 301 GTTAATATACACTCACAAGGAACCCAGTGCAGACGTGAAGACATCGACTCAAT 360
 QY 648 GAGCAGTTACCTGGGCTCTGAAAAGAAATTAGCGCCGAGTACCGCAACGACTGGAG 707
 DB 361 GAGCAGTTACCTGGGCTCTGAAAAGAAATTAGCGCGGAGTACCGCAACGACTGGAG 420
 QY 708 AAGGCGCTGCGGACCCAGTGTCTTACCTGGCGGAGAAATTGACACCGAGTACCCTTGC 767
 DB 421 AAGGCGCTGCGGACCCAGTGTCTTACCTGGCGGAGAAATTGACCTCGAGAGCCCTTGT 480
 QY 768 GCGCTTACACAGTACCACTGCTGGGAGGAGCACTAGCGCTCGGCAAGCTATGGTGGCG 827
 DB 481 GCGCTTACACAGTACCACTGCTGGGAGGAGCACTAGCGCTCGGCAAGCTATGGTGGCG 540
 QY 828 TTCTGCTTGGCTCTCTTCAACGCTGCTCTCAAGCGTCTGCGCGCGCCGCTCTAAGGAG 887
 DB 541 TTCTGCTTGGCTCTCTCTTCAACGCTGCTCTCAAGCGTCTGCGCGCGCCGCTCTAAG 600
 QY 888 CTGACACTGCTTCAACCGGAGCCTTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCT 947
 DB 601 CTGACACTGCTTCAACCGGAGCCTTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCT 660
 QY 948 TCTAAGTGCAGCTGCTGCGCGCTTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCT 990
 DB 661 TCTAAGTGCAGCTGCTGCGCGCTTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCT 703

RESULT 12
 A1826629 515 bp mRNA linear EST 21-DEC-1999
 LOCUS wk35e04.x1 NCI_CGAP_P22 Homo sapiens cDNA clone IMAGE:2417406 3'
 DEFINITION mRNA sequence.
 ACCESSION A1826629
 VERSION A1826629.1 GI:5447300
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 515)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bdtp/image/image.html
 Insert Length: 910 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 467.
 Location/Qualifiers
 1..515
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2417406"
 /sex="male"
 /tissue_type="normal prostate"
 /lab_host="DH10B"
 /clone_1lb="NCI CGAP P22"
 /note="Organ: prostate; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adapters (Pharmacia), digested with Not

[illegible]

```

VERSION      AT191844.1  GI:5339486
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homiidae; Homo.
REFERENCE    1 (bases 1 to 505)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Other ESTs: nk08a11.s1
             Contact: Robert Strausberg, Ph.D.
             Email: cgaps-remail.nih.gov
             Tissue Procurement: J. Jeffrey Medeiros, M.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             CDNA Library Preparation: Stratagene, Inc., David B. Kitzman,
             Ph.D.
             CDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bdnp/image/image.html

This read is a RESSEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
Information
This 5' resequenced clone has no previous 5' data to verify this
new read against
Putative full length read
Insert Length: 741 Std Error: 0.00
Seq primer: -40RP from Glbco
High quality sequence stop: 418.
Location/Qualifiers
1..505
/organism="Homo sapiens"
/mo_type="mRNA"
/cd_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab host="SOLR (kanamycin resistant)"
/clone_id="NCI CGAP C02"
/notes="Organ: colon; Vector: Bluescript SK-; Site_1:
Ecotri; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Bulk colon villous adenoma. 5' adaptor sequence:
5' GAATTCGCGACGAG 3' 3' adaptor sequence: 5'
CTCAGATTGTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."

ORIGIN
Query Match      34.4%; Score 489.2; DB 1; Length 505;
Best Local Similarity 99.2%; Pred. No. 2.5e-114;
Matches 502; Conservative 0; Mismatches 3; Indels 1; Gaps 1,

744  AAGTTCAACCGAATAGCCCTTGGGCGCTGACACACGTAACAACCTGGCGGACACTAC 803
Db    1  AAGTTCAACCGAATAGCCCTTGGGCGCTGACACACGTAACAACCTGGCGGACACTAC 60

804  GCTTCGCGCAAGCTATGGATGGAGCTTGTGCTTGTGCTCTCTCCAAAGTGTGCTTC 863
Db    61  GCTTCGCGCAAGCTATGGATGGAGCTTGTGCTTGTGCTCTCTCCAAAGTGTGCTTC 120

864  ACGCGCGCCCGCTCTACGAGGCGCTGGCACTGCTGACACCGGAGCTTGGCGCTTC 923
Db    121  ACGCGCGCCCGCTCTACGAGGCGCTGGCACTGCTGACACCGGAGCTTGGCGCTTC 180

924  GGGGCTTTCGCTTGGCTCCATCTTAAGCGATGCGCGCTCGCGCTCGGCTAGGCTCC 983
Db    181  GGGGCTTTCGCTTGGCTCCATCTTAAGCGATGCGCGCTCGCGCTCGGCTAGGCTCC 240

984  TCCGGCTCAACCACTAGTACGAGCGCGCTTTCGGATCAGGCTGGCAACCGGCGTCTG 1043

```

Db 241 TCGGCGCTGACCACTGATGAGGCGCGCTTCTGAGTCAAGCTGGCAACCGGCTCTG 300
QY 1044 TGCCTCTTCTCGGAGGGGCGTGTGAGTCTTCCAGTATGTTGGCCAGCGCTCTTGGC 1103
Db 301 TGCCTCTTCTCGGAGGGGCGTGTGAGTCTTCCAGTATGTTGGCCAGCGCTCTTGGC 360
QY 1104 ACCCTTCTGAGCAAGCGGCAAGGACTGAGCAAGGAGAGGGGGCTCACCCTTATC 1163
Db 361 ACCCTTCTGAGCAAGCGGCAAGGACTGAGCAAGGAGAGGGGGCTCACCCTTATC 420
QY 1164 CTGGGAGACCACTGACAGAGAGCGCTCTCCAGACTTAAATGTATCACCACTAAC 1223
Db 421 CTGGGAGACCACTGACAGAGAGCGCTCTCCAGACTT-AAATGTATCACCACTAAC 479
QY 1224 CTGTAGGGGGACCAATCTGACTC 1249
Db 480 CTGTAGGGGGACCAATCTGACTC 505

Search completed: January 22, 2006, 02:34:11
Job time : 6255 secs

THIS PAGE BLANK (USPTO)

QY	587	CATTAAATTTAACTCAACAGGGGACCCCGAGTGCATAGCTGAACGAGACCATTTGATCAAA	646
Db	321	AGTCAACATCACTCAAGGGGACCCCGTGCAGAGCTGAATGAGACCATCAATTAATTA	380
QY	647	CGAGCAGTTCACTTGGCGCTCTGAAAGAGAAATTACCGCCGCGAGTACAGCGAAAGCACTGGA	706
Db	381	CGAGGAGTTCACTGGCGCCTGGTGAAGAACTATCTGAGGAGTGTGCAAAAGGCTTTGGA	440
QY	707	GAAAGGGGCTGCCGAGCCCAAGTGTCTTACTTGGCGGAGAAAGTTCAACCGAGTACGCTTG	766
Db	441	GAAAGGGGCTGCCGAGCCCTGTGTACTTACTGAGAAAGTTCACTTCAAGAAAGCCCATG	500
QY	767	CGGCGCTGTACCAACAGTACCACTTGCGGGGACACTACGCTTGGCGACAGCTATGGGTGGC	826
Db	501	TGGCGCTTATACCGCCAGTACCGGCTTGGCGGGACACTACACTGACCAATGCTATGGGTGGC	560
QY	827	GTTCTGCTTCTGGGCTCTCTTCCAAAGTGTGCTCTTCAAGCCCGGCGCCGCTTACGAGAG	886
Db	561	ATTCTCTGTGGCTGTGGCTGGCCAAATGATGTCTTCAATGCTGTGATATATATGGTG	620
QY	887	CCTGGCACCTGGTGAACCAACCGAGCCTTGGCGCTCTTGCG---GGGTCTGGCTTGGGCTC	943
Db	621	CTAACATGCTATTTGGTCAAGGGGATTTTCAAGTGTGTGCTGTGCTTCTTCTTCATGGC	680
QY	944	CATCTCTAGCGTGGCGCTCTTGGCGCGCTCGGCTTGAAGCTCTTCCGCGCTACACACTCAAGTA	1003
Db	681	CACATCACTCAACCTCAACCCGTGCCCCGTGAACCTGGGCGCTTGTGCTGACATCACTCAACA	740
QY	1004	CGGCGCGCGCTTCTTGGGTGTCAGCTTGGCAACCGGCGCTCTGTGCGCTTCTTCTCGAAGGGGC	1063
Db	741	TGGCGCTCTCTTGTGATCAATTATCAACAGAGACTGTGTGTGTGCTGTGGGCTTGGC	800
QY	1064	CGTGTGATGCTTCCAGTATGTTGGCGCCAGCGGCTTCTGCAACCCCTTGAACCAAGGCGC	1123
Db	801	TATGGCGGTGGCGCCACAGATGACAGGCTTCAAGGCTTCTTCAACAGAGTGT	860
QY	1124	CAGGA 1129	
Db	861	GGATGA 866	

Query Match	25.0%	Score 354.8	DB 3	Length 2133
Best Local Similarity	63.8%	Pred. No. 9.3e-78		
Matches	555	Conservative	0	Mismatches 311; Indels 3; Gaps 1;
Dy	263	CAGCATGACCCCTGTGGAAAGGGGTAAGTACCTCTTTTAAACCCCGAGCCCGCGGATGCGCAGG	322	
Db	66	CAAGTGGCTTACTTTGGGACACATTTCCCTTCTATCTGTGACCCCGAGCCCAACCTTCC	125	

OY	323	CTTCAGCGTTCACATCGTCTCATTCGTTATTCTAAGTGTGTTTTGGCTCTAGACAGCAAGCTTCT	382
Db	126	GATGAGACCAACCTTTGGCAGACATCATCATATATCTTCTGACTGACCTGGCCACGTTAT	185
OY	383	GCTCATCTTGGCCGGAGATCCCGTGGCCACTCGCGCTGTGTTTTGGTGTGAGAGTTCCT	442
Db	186	CGTATCTCGCTGCTGGCANTTCGGGGAGAAACAGAGCTGTTCGGCTGCTTCGGGTGTAC	245
OY	443	CAGTCTGTTCAATAGCCGCAAAATTTGAGCTGTGCACTTCAGTGCAGAAATGGTTCGTGG	502
Db	246	CAGCTTATTCATCGGGGCTGCAAATCTGGCTGTGAATTCAGTTCTGAGTGTCTGTGGG	305
OY	503	TACAGTGAACCAACACATCTTCAAAAGCCTTACAGCGCAGCGCGGTACAGCCCGGT	562
Db	306	CCAGGTACGACCAACACATCATATCAAGGCCCTTTCAGTTCTGATGATCAGCGCTGATAT	365
OY	563	CGGTCTGCTCGTGGGCGCTGAGAGGCACTTAATATTACACTCAACAGGACCCCAAGTACCA	622
Db	366	TGGCTGTAGAGTGGGGGTGGGTGAGTGCACATCACACTCAACAGGACCCCGTGGACCA	425
OY	623	GCTGAACGACACCATTTGACTACACAGACAGTTCACTGGCGTCTGAAAGGAATTACG	682
Db	426	GCTAAATGAGCACTATCAATTACAACAGAGAGTTCACTGGGCGCTGGGTGGAATTAAGC	485
OY	683	CGCGAGATACCGCAACGCACTGAGAGAGGGGCTGCCGAGCCAGTGTCTACCTGGCGGA	742
Db	486	TGAGGAGTGTGCANAGGCTCTGAGAGAGGGGCTGCCAGCCCTGTGTTGTACTAGCTGA	545
OY	743	GAAGTTCACACCGAGTAGCCCTTCGGCGCTGTACCAACCACTACCACTTGGCGGACACTA	802
Db	546	GAAGTTCACCTCAAGAAAGCCCATGTGGCTGTATACCGCCAGTACCGGCTGGCGGACACTA	605
OY	803	CGCCTCGGCCACGCTAATAGGTGAGGGTCTGTGTTCTGGGTCTCTTCCACAGTGTGCTTC	862
Db	606	CACCTCAGCCATGTATATGAGTGGCAATCTCTGTCTGGGTGTGGCCAAATGTGATGCTTC	665
OY	863	CACGCGGCGCCGCTCTACACGAGGCTTGGCACTGTGACCAACCGAGGCTTTCGCGCTCTT	922
Db	666	CATGCTGTGTGTGATATATGTTGGCTACATGCTATTTGGCCACGGGCAATCTTTCAGGTGTT	725
OY	923	CG---GGGTCTTGGCCTTGGGCTTCATCTCTAAGGTGCGGCTCTGCGCGCTCGCTAAG	979
Db	726	GGCTCTGCTCTTCTTCTTCACATGAGCAACATCACTCACTCAACCTGTCCCTCGACCTGGG	785
OY	980	CTCCTCGCGGCTACCACTGATACGGGCGCGGCTTCTGGGTCAACGCTGGCAACCGGCGT	1039
Db	786	CGCTTCTGTGTGCTGATACTACACATGGGCGCTGCTTCTGATTCACATTTGACCAAGACT	845
OY	1040	CTTGTGCTCTTCTCTCGAGAGGGGCGTGTGAGTCTCCAGTATGTTGGGCCCAAGCTCT	1099
Db	846	GCTGTGTGTGCTGTGGGCGCTGTATGAGCGGTGAGCCACAGATGACGCTCACAGGCT	905
OY	1100	TCGCACCCCTTTCGACCCAAAGCGCCAAAGA	1129
Db	906	GAAAGCTTTCTTCAACCAAGATGTGAGTGA	935

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 223 GGGAGCGCCGAGGTGTCGCGAGTCCGCGCGCTGCGACTGACGGCTTCTTC 282
Qy 786 CACCTGGCGGAGACATACGCTTGGCCAGCTATGGGTGGCTTCTGCTTCCCTC 845
Db 283 CAGGTGGTGGGACGAGATGACGCGCGCTGACGCGAGGCCACCGCTGATGAC 342
Qy 846 TCCAAAGTGTGCTTCCACGCGCGCGCTTACGAGAGCTTGGCACTGATGAC 905
Db 343 GCTTCTTACGCTGCGCTCCCGCGACGACGCGCGCGCGCGCGCGCGCGCG 402
Qy 906 GAGACCTTGGCGCTTCTTGGGAGTCTTCCCTTGGCTTCC 945
Db 403 TGGGCTTACGCGCGAGCTTACGCGCGCGCTTCCGCTTCA 442

RESULT 9
US-09-252-991A-11029/c
; Sequence 11029, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11029
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11029

Query Match 3.3%; Score 47.6; DB 3; Length 420;
Best Local Similarity 48.2%; Pred. No. 0.035;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 793 CGGAGACATACGCTTGGCGCGCGCTTATGGGTGGCTTCTGCTTCTTCCACG 852
Db 382 CGGTGGAACGCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 323
Qy 853 TGTGCTCTTCCAGCGCGCGCGCTTATGCGAGGCTTGGCACTGCTGACCA 912
Db 322 CGGTGACGAGAGCGCGCGGTGTTCAAGAGCTGACCGGTCTTGGCGATG 263
Qy 913 TCGCGCTTCTGAGGCTTCTGCTTGGCTTCCATCTTAAAGCTGCGCTT 972
Db 262 TCTCCCTGAGCGCTTGGGACCTTCTGCTGCTGCTGCTGCTGCTGCTG 203
Qy 973 GCTTACGCTTCTGCGCTTACCACTGACGCGCGCTTGGGTGACGCTG 1032
Db 202 CTTTGGCTTCCAGCGCGCGGTGTTCACTTCTGCTTCTGCTGCTGCTG 143
Qy 1033 CCGGCGCTTGTGCTTCTTCTTCCGAGAGGCGCTG 1070
Db 142 GCGGCTTGGCTTACCTTCTTCTGCTTCTGCTGCGCGCGCGCG 105

RESULT 10
US-09-252-991A-11226
; Sequence 11226, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11226
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11226

Query Match 3.3%; Score 47.6; DB 3; Length 1683;
Best Local Similarity 48.2%; Pred. No. 0.049;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 793 CGGAGACATACGCTTGGCGCGCGCTTATGGGTGGCTTCTGCTTCTTCCACG 852
Db 864 CGGTGGAACGCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
Qy 853 TGTGCTCTTCCAGCGCGCGCGCTTATGCGAGGCTTGGCACTGCTGACCA 912
Db 924 CGGTGACGAGAGCGCGCGGTGTTCAAGAGCTGACCGGTCTTGGCGATG 983
Qy 913 TCGCGCTTCTGAGGCTTCTGCTTGGCTTCCATCTTAAAGCTGCGCTT 972
Db 984 TCTCCCTGAGCGCTTGGGACCTTCTGCTGCTGCTGCTGCTGCTGCTG 1043
Qy 973 GCTTACGCTTCTGCGCTTACCACTGACGCGCGCTTGGGTGACGCTG 1032
Db 1044 CTTTGGCTTCCAGCGCGCGGTGTTCACTTCTGCTTCTGCTGCTGCTG 1103
Qy 1033 CCGGCGCTTGTGCTTCTTCTTCCGAGAGGCGCTG 1070
Db 1104 GCGGCTTGGCTTACCTTCTTCTGCTTCTGCTGCGCGCGCG 1141

RESULT 11
US-09-252-991A-11111
; Sequence 11111, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11111
; LENGTH: 1974
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11111

Query Match 3.3%; Score 47.6; DB 3; Length 1974;
Best Local Similarity 48.2%; Pred. No. 0.051;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 793 CGGAGACATACGCTTGGCGCGCGCTTATGGGTGGCTTCTGCTTCTTCCACG 852
Db 731 CGGTGGAACGCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790
Qy 853 TGTGCTCTTCCAGCGCGCGCGCTTATGCGAGGCTTGGCACTGCTGACCA 912
Db 791 CGGTGACGAGAGCGCGCGGTGTTCAAGAGCTGACCGGTCTTGGCGATG 850
Qy 913 TCGCGCTTCTGAGGCTTCTGCTTGGCTTCCATCTTAAAGCTGCGCTT 972
Db 851 TCTCCCTGAGCGCTTGGGACCTTCTGCTGCTGCTGCTGCTGCTGCTG 910

QY	973	GCTTAGGCTCTCCGCGCTCAACA	CTAGTAAGGAGCGCCCTTCCAGGTCACGCTGGCAA	1032
Db	911	CTTTGCGCTCCGACCCGAGCGCGGGGGTTCATCTCGCCTCTCTGCTGCTGGTGGTCG	970	
QY	1033	CCGGGCGTCTGTGCTCTTCTCTCGAGAGGGCGCGTGGTG	1070	
Db	971	GCGGCTGCTACCCCTGTTGCGCCCTGACCGCGACCGGTG	1008	

RESULT 12

```

US-09-252-991A-10956/c
; Sequence 10956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10956
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10956

```

Query Match	Score	DB	Length
3.3%	47.6	3	2283

QY	793	CGGAGCACTAGCCCTCGGACAGCGTATAGGAGGCGTTTGCTTTCGCTCCTTCGAACG	852
Db	2081	CGGTGGAAAAAGCTTCTCTTCAATGCCCTGCTGCTGCTGCGCATCTGCACTATATCCACTTCGCTGG	20222
QY	853	TGCTGCTTCCACGCGCGGCGCCCGCTTACGGAAGGCTGGACATGCTGACCCAGCGGAGCTT	912
Db	2021	CGGTGACCGAGAAACGCGCGCGCTGTATCAAGACTGGAACGCTGCTTCTTGCGAGATCGCCGCTT	19628
QY	913	TGCGGCTCTTTCGGGGTCTTGGCGCTTGGCGTCAATCTCTAAGAGTGCCTGTCGCCCGCTCC	972
Db	1961	TCTCCCTGAGCTGCTGTGGGCAACCTTCTCTGTCGCTTCGGGAGGTGTACACTTCGGTGACG	19022
QY	973	GCTTAGGCTCTTCGCGCGTCAACAATCAATGTAAGGAGCCGCTTCTGAGTCAACGCTGGCAA	10322
Db	1901	CTTTCGCGCTCGAACCCGGAGCGCGGGAGTTCAATCTGACCTTCTGCTGTGCTGGTGGTTCG	18422
QY	1033	CGGCGGTCCTGTGCTCTTCTCTCGAAGAGGCGCGTGGTG	1070
Db	1841	GGGGCTCGCTGACCTCTGTTCGCCCTTCGCGCGGCGCGGTG	1804

RESULT 13

```

US-09-614-912-49
; Sequence 492, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cairni, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes

```

FILE REFERENCE: B81378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912.2
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97
SEQ ID NO: 49

Query Match	3.3%	Score 46.4	DB 3	Length 1431
-------------	------	------------	------	-------------

QY	545	GTATACAGCCCGATGTGGATCTGTCTGTGGGCGCTGGAGGGCAATTAATATATACATCAACAGGG	608
Db	33	GCTTCTCCTCTCTCTCTCTCAACATGATGTGGCTGGACGATATCCCCCTCCCTTTG	92
QY	609	ACCCGATGTCATCACTGTAACGAGACATTGACTTACACAGACGATTCACCTGGCGCTTG	668
Db	93	TTTGACGCTGCCCGCTGAGGGCGCTTCGCAATATCCGACGAGTTCAATCTGGCGGGCG	152
QY	669	AAAGAGATTAACGCCCGCGAGTACGCGAACCACTGGAGAAAGGGGCTGCCGACCAATG	728
Db	153	GACGAGAGCCCCACCCCGAATCCGCGCCAGAGAGCTGGCGCTGGCTCATGACCTCTCC	212
QY	729	CTCTACCTGGCGGAGAGTTCAACCG--AGTAGCCCTTGGCGGCTGTACACAGATAC	785
Db	213	GGGGAAGCCCGCGAATGATGATCCGGAGGATCCGGCGCGCTGACACTGACAGGTTCTTC	272
QY	786	CACCTGACGGGACACTACGCTTCGCGCACGCTATAGGGTGGCGTTCTCTTCTGACTCTC	845
Db	273	CAGTGTGTGGGCAAGGATGACGCGGCGCTGACGCGGAGGCCACCGCTGATGAGAC	332
QY	846	TCCAACTGTCTCTTCCACGCGCGGCCCGCTTACGAGGCGCTGGCACTGTGACCAAC	905
Db	333	GCTTCTTACGCTCCGCTCCCGACAAAGCAGCGCGCAGCGCCGCTCAGGGGGACAGC	392
QY	906	GGAGCCTTCGCGCTTTCGGGGTCTTGCCTTTGGGCTTCA	945
Db	393	TGCGGCTACGCGACGACTTACAGGCGCGGTTCCGCTTCA	432

RESULT 14

US-09-949-016-88511
; Sequence 88511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 88511
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-88511

Query Match 3.2%; Score 46; DB 3; Length 601;
Best Local Similarity 57.7%; Pred. No. 0.095;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1035 GGGGCTGTCGCTCTCTCTCGAGGGGCGGTGAGTCTCCAGTATGTTCCGCCAGC 1094
DB 63 GGGGCTGTCGCTCTCTCGAGGGGCGGTGAGTCTCCAGTATGTTCCGCCAGC 122
QY 1095 GCTCTTCGACCCCTTCTGACCAAGCGCCAGACTGCAGCCAGAGAGAGGGGGCTCA 1154
DB 123 AGGGTCTCAGCCCTGCGGCCCAATGATAGTACACTGAGGCCGCGAGATAGGGGGCTTT 182
QY 1155 CCTCTTATCTCGCGCACCCAC 1176
DB 183 AGAGTAGCCGCTGGGAGCCTC 204

RESULT 15
US-09-949-016-14248/c
Sequence 14248, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14248
LENGTH: 8991
TYPE: DNA
ORGANISM: Human
US-09-949-016-14248

Query Match 3.2%; Score 46; DB 3; Length 8991;
Best Local Similarity 57.7%; Pred. No. 0.19;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1035 GGGGCTGTCGCTCTCTCTCGAGGGGCGGTGAGTCTCCAGTATGTTCCGCCAGC 1094
DB 3532 GGGGCTGTCGCTCTCTCGAGGGGCGGTGAGTCTCCAGTATGTTCCGCCAGC 3473
QY 1095 GCTCTTCGACCCCTTCTGACCAAGCGCCAGACTGCAGCCAGAGAGAGGGGGCTCA 1154
DB 3472 AGGGTCTCAGCCCTGCGGCCCAATGATAGTACACTGAGGCCGCGAGATAGGGGGCTTT 3413

QY 1155 CCTCTTATCTCGCGCACCCAC 1176
DB 3412 AGAGTAGCCGCTGGGAGCCTC 3391

Search completed: January 22, 2006, 02:40:03
Job time : 288 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 22:28:47 ; Search time 1266 Seconds
(without alignments)
9137.470 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaagtaacgctc.....ctttttcttttctttt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.Main:*
1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/us08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/us10_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/us10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/us10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/us10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/us10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/us10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1421	100.0	1421	US-10-187-657-2	Sequence 2, Appl1
2	1421	100.0	1450	US-10-643-795A-5	Sequence 5, Appl1
3	1421	100.0	1450	US-10-948-518-5	Sequence 54, Appl1
4	1418.6	99.8	1474	US-10-264-237-1097	Sequence 1097, Ap
5	1415.8	99.6	1420	US-10-187-657-4	Sequence 4, Appl1
6	1412	99.4	1491	US-10-773-236-187	Sequence 187, App
7	693	48.8	1594	US-10-187-657-9	Sequence 9, Appl1
8	459.6	32.3	522	US-10-187-657-5	Sequence 69, Appl1
9	413	29.3	580	US-10-773-236-69	Sequence 2367, Ap
10	402.8	28.3	406	US-10-450-763-3367	Sequence 72673, A
11	396.6	27.9	1376	US-09-925-065A-72673	Sequence 1195, Ap
12	391	27.5	2684	US-10-094-749-1195	Sequence 7, Appl1
13	385.8	27.1	450	US-10-187-657-7	Sequence 424, App
14	354.8	25.0	1029	US-09-759-130B-424	Sequence 424, Appl
15	354.8	25.0	1029	US-10-042-431-54	Sequence 54, Appl1
16	354.8	25.0	1029	US-10-741-790-424	Sequence 424, App
17	354.8	25.0	2133	US-09-759-130B-423	Sequence 423, App
18	354.8	25.0	2133	US-10-042-431-53	Sequence 53, Appl1
19	354.8	25.0	2133	US-10-741-790-423	Sequence 54, App
20	322	22.7	506	US-10-187-657-6	Sequence 6, Appl1
21	300	21.1	346	US-10-187-657-3	Sequence 3, Appl1
22	245.8	17.3	2684	US-10-094-749-1195	Sequence 1195, Ap
23	217.4	15.3	395	US-09-925-065A-100165	Sequence 100165, Ap

24	217.4	15.3	395	4	US-09-925-065A-145294	Sequence 145294, A
25	200.6	14.1	439	5	US-10-027-632-91970	Sequence 91970, A
26	200.6	14.1	439	6	US-10-027-632-91970	Sequence 91970, A
27	199.4	14.0	232	3	US-09-783-590-9482	Sequence 9482, Ap
28	190	13.4	277	4	US-09-925-065A-100166	Sequence 100166, A
29	170	12.0	250	5	US-10-187-657-8	Sequence 8, Appl1
30	161.8	11.4	1338	9	US-10-450-763-3511	Sequence 3511, Ap
31	65.6	4.6	1602	10	US-11-097-143-29060	Sequence 29060, A
32	65.6	4.6	1808	10	US-11-097-143-8903	Sequence 8903, Ap
33	64	4.5	512	5	US-10-106-698-3277	Sequence 3277, Ap
34	55.2	3.9	1783	7	US-10-437-963-340	Sequence 340, App
35	54.8	3.9	555	7	US-10-767-701-13710	Sequence 13710, A
36	53	3.7	3024	7	US-10-437-963-10408	Sequence 10408, A
37	52.6	3.7	1087	8	US-10-425-115-140838	Sequence 140838, A
38	52	3.7	8134	7	US-10-437-963-93461	Sequence 93461, A
39	51.6	3.6	1280	7	US-10-437-963-35623	Sequence 35623, A
40	51.6	3.6	1506	7	US-10-437-963-40658	Sequence 40658, A
41	51.2	3.6	1123	7	US-10-425-114-15938	Sequence 15938, A
42	51.2	3.6	1454	8	US-10-425-115-46357	Sequence 46357, A
43	50.8	3.6	2508	7	US-10-437-963-2740	Sequence 2740, Ap
44	50.6	3.6	1455	6	US-10-156-761-5889	Sequence 5889, Ap
45	50.6	3.6	9025608	6	US-10-156-761-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-187-657-2
Sequence 2, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Laasek, Amy K.W.
APPLICANT: Baughn, Mariah R.
APPLICANT: Aaimai, Valda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT APPLICATION NUMBER: US/10/187,657
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: PCT/US00/07817
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 1421
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20030068311A1 4901066CBI
US-10-187-657-2
Query Match 100.0%; Score 1421; DB 5; Length 1421;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAACCAAGAAAGTAACGCTTACAGACAGTGAAGAAATGTTGCTGCGCCGGCTAGAAAA 60
1 GAACCAAGAAAGTAACGCTTACAGACAGTGAAGAAATGTTGCTGCGCCGGCTAGAAAA 60
1 GAACCAAGAAAGTAACGCTTACAGACAGTGAAGAAATGTTGCTGCGCCGGCTAGAAAA 60
61 CTCTGTGCTTACCAACCCGAGGCTTGAAGACAGCCCACTTCAAGCTTCTTAACGGA 120
61 CTCTGTGCTTACCAACCCGAGGCTTGAAGACAGCCCACTTCAAGCTTCTTAACGGA 120
121 GAGGTGAGAGACTGAGCTTACCAAGCCCACTGCTGCTTGAAGAGAGAGAGAGAGAG 180
121 GAGGTGAGAGACTGAGCTTACCAAGCCCACTGCTGCTTGAAGAGAGAGAGAGAGAG 180
181 CCAAG 240
181 CCAAG 240

QY 241 CGTGCAGCACTCGGCGCGCGGTGACATGACCTGTTGGAACGGGCTATGCGCTTTTACC 300
DB 241 CGTGCAGCACTCGGCGCGCGGTGACATGACCTGTTGGAACGGGCTATGCGCTTTTACC 300
QY 301 CCCAGCCCCCGGATGCGCGACAGGCTTACGCTTCACTGCTCATGTTATTTAGTGT 360
DB 301 CCCAGCCCCCGGATGCGCGACAGGCTTACGCTTCACTGCTCATGTTATTTAGTGT 360
QY 361 TGGCTCTAGACAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCACTCGCGCTG 420
DB 361 TGGCTCTAGACAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCACTCGCGCTG 420
QY 421 TTTGGTTGGTGAAGTCTTCTGCTCATGCTGTTTCAATAGCGGAAATTTGGCTGTGACT 480
DB 421 TTTGGTTGGTGAAGTCTTCTGCTCATGCTGTTTCAATAGCGGAAATTTGGCTGTGACT 480
QY 481 TCAGTGAGAAATGTTGTTGTTGTTACATGTAACCAACACATCTCTCAAGCTTCAAGC 540
DB 481 TCAGTGAGAAATGTTGTTGTTGTTACATGTAACCAACACATCTCTCAAGCTTCAAGC 540
QY 541 CACCGCGCTTACAGCCCGTGTGGTGTCTGCTGGGCGCTGAGGGCATTAATATTAAC 600
DB 541 CACCGCGCTTACAGCCCGTGTGGTGTCTGCTGGGCGCTGAGGGCATTAATATTAAC 600
QY 601 TCAAGGGAACCCAGTGCATGCTGTAACGAGACATTAATGTAACAAGAGAGTCACT 660
DB 601 TCAAGGGAACCCAGTGCATGCTGTAACGAGACATTAATGTAACAAGAGAGTCACT 660
QY 661 GGGCTCTGAAGAAATTAAGCGCGGAGTACGCGAAACGCACTGGAAGAGGCGTCCG 720
DB 661 GGGCTCTGAAGAAATTAAGCGCGGAGTACGCGAAACGCACTGGAAGAGGCGTCCG 720
QY 721 ACCAGTGTCTACCTGCGGGAAGTTACACCGAGTACCGCTTGGCTGTACAC 780
DB 721 ACCAGTGTCTACCTGCGGGAAGTTACACCGAGTACCGCTTGGCTGTACAC 780
QY 781 AGTACCACTGCGGGAACATGACGCTCGGACAGCTATGGTGGCTTCTGCTTGGC 840
DB 781 AGTACCACTGCGGGAACATGACGCTCGGACAGCTATGGTGGCTTCTGCTTGGC 840
QY 841 TCTCTTCAAGTGTGCTCTTCAACCGCGGCGCTTCAAGAGGCTTGGCACTGTA 900
DB 841 TCTCTTCAAGTGTGCTCTTCAACCGCGGCGCTTCAAGAGGCTTGGCACTGTA 900
QY 901 CCAAGGAGCTTGGGCTTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 960
DB 901 CCAAGGAGCTTGGGCTTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 960
QY 961 TCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1020
DB 961 TCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1020
QY 1021 TCAAGCTGGAACCGGCTTCTGCTTCTTGGAGGGGCGCTGTTGAGTCTCAAGT 1080
DB 1021 TCAAGCTGGAACCGGCTTCTGCTTCTTGGAGGGGCGCTGTTGAGTCTCAAGT 1080
QY 1081 ATGTTGGGCGGAGGCTTCTGCAACCTTCTGGAACCAAGCGGCTTGGAGGCAAG 1140
DB 1081 ATGTTGGGCGGAGGCTTCTGCAACCTTCTGGAACCAAGCGGCTTGGAGGCAAG 1140
QY 1141 AGAGAGGGGCTTCACTTATCTGCGGACCACTGCAACAAGGCGCTTCCAG 1200
DB 1141 AGAGAGGGGCTTCACTTATCTGCGGACCACTGCAACAAGGCGCTTCCAG 1200
QY 1201 ACTTAAATGTAACCACTTATCTGAGGGGAGACCAATCTGACTCTTCCGCT 1260
DB 1201 ACTTAAATGTAACCACTTATCTGAGGGGAGACCAATCTGACTCTTCCGCT 1260
QY 1261 TGGGACATGCGAGGCGGGAAGAGAGTGGCGGCAAGGCTGGGCGAGAGAGTCCAGAA 1320
DB 1261 TGGGACATGCGAGGCGGGAAGAGAGTGGCGGCAAGGCTGGGCGAGAGAGTCCAGAA 1320

QY 1321 GGGCACTGAGCGCTGTGGCGGAGGCTTGGAGATCCCGAGGACCAAGGAAAGTCTCC 1380
DB 1321 GGGCACTGAGCGCTGTGGCGGAGGCTTGGAGATCCCGAGGACCAAGGAAAGTCTCC 1380
QY 1381 TGGGCGGATCTGTAATAAATTAACCTTTTCTTTTGT 1421
DB 1381 TGGGCGGATCTGTAATAAATTAACCTTTTCTTTTGT 1421

RESULT 2

US-10-643-795A-5
; Sequence 5, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN PRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/643,795A
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 5
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: Unsure
; LOCATION: 2
; OTHER INFORMATION: Unknown base
US-10-643-795A-5

Query Match 100.0%; Score 1421, DB 8; Length 1450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGAAAGTAACGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGTAAAGAAA 60
DB 9 GAACAGAAAGTAACGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGTAAAGAAA 68
QY 61 CTCTGTGGTACCAACCCCAAGGCGTTGAGAGAGAGCCCACTTCAAGCTTCTTAAAG 120
DB 69 CTCTGTGGTACCAACCCCAAGGCGTTGAGAGAGAGCCCACTTCAAGCTTCTTAAAG 128
QY 121 GAGGTGAGAGACTGAGACTTCAAGAGCCCACTTGGTCCAGGCTTGTACGCAAGAGAG 180
DB 129 GAGGTGAGAGACTGAGACTTCAAGAGCCCACTTGGTCCAGGCTTGTACGCAAGAGAG 188
QY 181 CCAAGAGAGCGCTTCCGCGCTCAAGAGAGCCCAAGTGTGCTGCTTCCGCGCTG 240
DB 189 CCAAGAGAGCGCTTCCGCGCTCAAGAGAGCCCAAGTGTGCTGCTTCCGCGCTG 248
QY 241 CGTGCAGCACTCGGCGGCGTGCAGATGACCTTGTGAACGGGCTACTGCTTTTACC 300
DB 241 CGTGCAGCACTCGGCGGCGTGCAGATGACCTTGTGAACGGGCTACTGCTTTTACC 300

QY 301 CCCAGCCCGGCGATGCGCGAGGCTTCAAGGCTCACTGCTCATTCGTTATTTCTAGTGT 360
 DB 309 CCCAGCCCGGCGATGCGCGAGGCTTCAAGGCTCACTGCTCATTCGTTATTTCTAGTGT 368
 QY 361 TGGCTCTAGCAGCAAGCTTCTGCTCATCTTGCCTGCGGAGTCCGTGAGCACTCGCGCTG 420
 DB 369 TGGCTCTAGCAGCAAGCTTCTGCTCATCTTGCCTGCGGAGTCCGTGAGCACTCGCGCTG 428
 QY 421 TTTGGTTGATGAGTCTTCTCAAGTCTGTTCAATAGGCGAGAAATTTGGCTGTGCACT 480
 DB 429 TTTGGTTGATGAGTCTTCTCAAGTCTGTTCAATAGGCGAGAAATTTGGCTGTGCACT 488
 QY 481 TCAGTGCAGAAATGTTGCTGGGATGAGTGAACCAACATCTTCAAAAGCTTCAGG 540
 DB 489 TCAGTGCAGAAATGTTGCTGGGATGAGTGAACCAACATCTTCAAAAGCTTCAGG 548
 QY 541 CAGCGCGCTTACAGCCCGTGTGCTGTCTGTGGGCTTGAAGGCAATTAATTATAC 600
 DB 549 CAGCGCGCTTACAGCCCGTGTGCTGTCTGTGGGCTTGAAGGCAATTAATTATAC 608
 QY 601 TCACAGGGAACCCAGTGCATCACTGAAGACCACTTGACTTCAACAGAGCTTCACT 660
 DB 609 TCACAGGGAACCCAGTGCATCACTGAAGACCACTTGACTTCAACAGAGCTTCACT 668
 QY 661 GGGGCTGAAAGAAATTAAGCGCGGAGATGACGCAACGCACTGGAAGAGGGCTGCGG 720
 DB 669 GGGGCTGAAAGAAATTAAGCGCGGAGATGACGCAACGCACTGGAAGAGGGCTGCGG 728
 QY 721 ACCCAAGTGTCTACCTGCGGAGAAAGTTCAACCGAGTACGCTTGGGCTGTACAC 780
 DB 729 ACCCAAGTGTCTACCTGCGGAGAAAGTTCAACCGAGTACGCTTGGGCTGTACAC 788
 QY 781 AGTACCACTGCGGAGACACTAGGCTCGGCGACGCTAATGAGGTCTGTCTGTGCG 840
 DB 789 AGTACCACTGCGGAGACACTAGGCTCGGCGACGCTAATGAGGTCTGTCTGTGCG 848
 QY 841 TCTCTCCAAGTGTCTCTCCAAGCGGCGGCTCTACGGAAGGCTTGGCACTGCTA 900
 DB 849 TCTCTCCAAGTGTCTCTCCAAGCGGCGGCTCTACGGAAGGCTTGGCACTGCTA 908
 QY 901 CCAACCGAGCTTCTGCGGCTTCTGCGGAGTCTTGCCTTGCATCTTGAAGTGC 960
 DB 909 CCAACCGAGCTTCTGCGGCTTCTGCGGAGTCTTGCCTTGCATCTTGAAGTGC 968
 QY 961 TCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGG 1020
 DB 969 TCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGG 1028
 QY 1021 TCACGCTGCAACCGGCGTCTGTGCTCTTCTCGGAAGGCGCTGTGAGTCTCA 1080
 DB 1029 TCACGCTGCAACCGGCGTCTGTGCTCTTCTCGGAAGGCGCTGTGAGTCTCA 1088
 QY 1081 ATGTTGCGCCAGCGCTTCTGCGACCTTCTGGAACCAAGCGCAAGCACTGAGCCAG 1140
 DB 1089 ATGTTGCGCCAGCGCTTCTGCGACCTTCTGGAACCAAGCGCAAGCACTGAGCCAG 1148
 QY 1141 AAGAGGGGGGCTCACTCTTATCTGTGGGACCACTGCAAGAGGCGCTCTCCAG 1200
 DB 1149 AAGAGGGGGGCTCACTCTTATCTGTGGGACCACTGCAAGAGGCGCTCTCCAG 1208
 QY 1201 ACTTAATATGATCAACCACTGATGAGGGGACCAATCTGACTCTTCCCGCT 1260
 DB 1209 ACTTAATATGATCAACCACTGATGAGGGGACCAATCTGACTCTTCCCGCT 1268
 QY 1261 TGGGACATGCGAGGCGGGAAGAGTCCCGCAGGCTGAGGCAAGAGTCTCCAGAA 1320
 DB 1269 TGGGACATGCGAGGCGGGAAGAGTCCCGCAGGCTGAGGCAAGAGTCTCCAGAA 1328
 QY 1321 GGGCACTGAGCGCTGCTGCGGAGGCTTCAAGCACTCCGAGGCAACAGGAAAGTCTCC 1380
 DB 1329 GGGCACTGAGCGCTGCTGCGGAGGCTTCAAGCACTCCGAGGCAACAGGAAAGTCTCC 1388

QY 1381 TGGGGGATCTGTAATAAACCTTTTCTTTGTTTT 1421
 DB 1389 TGGGGGATCTGTAATAAACCTTTTCTTTGTTTT 1429
 RESULT 4
 US-10-264-237-1097
 / Sequence 1097, Application US/10264237
 / Publication No. US20040009491A1
 / GENERAL INFORMATION:
 / APPLICANT: Biese et al.
 / TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 / FILE REFERENCE: PA131P1
 / CURRENT APPLICATION NUMBER: US/10/264,237
 / PRIOR FILING DATE: 2002-10-04
 / PRIOR APPLICATION NUMBER: PCT/US01/16450
 / PRIOR FILING DATE: 2001-05-18
 / PRIOR APPLICATION NUMBER: US 60/205,515
 / NUMBER OF SEQ ID NOS: 2876
 / SOFTWARE: PatentIn Ver. 3.1
 / SEQ ID NO: 1097
 / LENGTH: 1474
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-10-264-237-1097
 Query Match 99.8%; Score 1418.6; DB 6; Length 1474;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1418; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAACGAGAAAGTAACGGCTACAGACAGTGAATAATGTTGCTGCGCGCTAGAAAA 60
 DB 19 GAACGAGAAAGTAACGGCTACAGACAGTGAATAATGTTGCTGCGCGCTAGAAAA 78
 QY 61 CTCTGTGCTACCAACCCGAGGCTTGAAGACAGCCCACTTCAAGCTTCTTAAACGA 120
 DB 79 CTCTGTGCTACCAACCCGAGGCTTGAAGACAGCCCACTTCAAGCTTCTTAAACGA 138
 QY 121 GAGGTGCGAGTCTGACTTCAACGAGCCCACTGCTGCGAGCTTGAAGAGAGAG 180
 DB 129 GAGGTGCGAGTCTGACTTCAACGAGCCCACTGCTGCGAGCTTGAAGAGAGAG 198
 QY 181 CCAAGAGCGGCTCTCCGCGCTTCAAGGAGCCCAAGTGTGCTGCTGCGCGCTG 240
 DB 199 TCAAGAGCGGCTCTCCGCGCTTCAAGGAGCCCAAGTGTGCTGCTGCGCGCTG 258
 QY 241 CGTGAAGCACTGCGCGGCTGCAAGATGACCTGTGGAAGCGGCTTCTGCTTAC 300
 DB 259 CGTGAAGCACTGCGCGGCTGCAAGATGACCTGTGGAAGCGGCTTCTGCTTAC 318
 QY 301 CCCAGCCCGGATGCGCGAGGCTTCAAGGCTCACTGCTCATTCGTTATTTAGTGT 360
 DB 319 CCCAGCCCGGATGCGCGAGGCTTCAAGGCTCACTGCTCATTCGTTATTTAGTGT 378
 QY 361 TGGCTCTAGCAGCAAGCTTCTGCTCATCTTGCCTGCGGAGTCCGTGAGCACTCGCGCTG 420
 DB 379 TGGCTCTAGCAGCAAGCTTCTGCTCATCTTGCCTGCGGAGTCCGTGAGCACTCGCGCTG 438
 QY 421 TTTGGTTGATGAGTCTTCTCAAGTCTGTTCAATAGGCGAGAAATTTGGCTGTGCACT 480
 DB 439 TTTGGTTGATGAGTCTTCTCAAGTCTGTTCAATAGGCGAGAAATTTGGCTGTGCACT 498
 QY 481 TCAGTGCAGAAATGTTGCTGGGATGAGTGAACCAACATCTTCAAAAGCTTCAGG 540
 DB 499 TCAGTGCAGAAATGTTGCTGGGATGAGTGAACCAACATCTTCAAAAGCTTCAGG 558
 QY 541 CAGCGCGCTTACAGCCCGTGTGCTGTCTGCTGTGAGGCTTGAAGGCAATTAATTATAC 600
 DB 559 CAGCGCGCTTACAGCCCGTGTGCTGTCTGCTGTGAGGCTTGAAGGCAATTAATTATAC 618
 QY 601 TCACAGGGAACCCAGTGCATCACTGAAGACCACTTGACTTCAACAGAGCTTCACT 660

```

Db      619 TCACAGGACCCGAGTGCATCAGTCGAAACGACATTTGTAACAAAGAGCTTCACT 678
Qy      661 GGGGTCTGAAAGAAATTACCGCGCGGAGTACGCGAACGACATGAGAGAGGGCTGCGG 720
Db      679 GGGGTCTGAAAGAAATTACCGCGCGGAGTACGCGAACGACATGAGAGAGGGCTGCGG 738
Qy      721 ACCCAGTGTCTACTCTGCGGAGAAAGTTCAACACCGAGTACCCCTTGGCGCTGTACCA 780
Db      739 ACCCAGTGTCTACTCTGCGGAGAAAGTTCAACACCGAGTACCCCTTGGCGCTGTACCA 798
Qy      781 AGTACCACTGCGGCGGAGCACTACGCTCGGCGACGCTATGAGTGGCTTCTGCTTGGC 840
Db      799 AGTACCACTGCGGCGGAGCACTACGCTCGGCGACGCTATGAGTGGCTTCTGCTTGGC 858
Qy      841 TCCCTCCAAAGTGTCTCTCCACGCGCGCGCTCTAAGAGGCTGTGCACTGTGTA 900
Db      859 TCCCTCCAAAGTGTCTCTCCACGCGCGCGCTCTAAGAGGCTGTGCACTGTGTA 918
Qy      901 CCAACGAGGCTTTCGCGCTCTTCGCGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 960
Db      919 CCAACGAGGCTTTCGCGCTCTTCGCGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 978
Qy      961 TCTGCGGCTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 1020
Db      979 TCTGCGGCTTCGCGCTTTCGCGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 1038
Qy      1021 TCACGCTGCGCAACGCGGCTCTGTGCTTTCGCGGAGGGCGGTGTAGTCTCCAGT 1080
Db      1039 TCACGCTGCGCAACGCGGCTCTGTGCTTTCGCGGAGGGCGGTGTAGTCTCCAGT 1098
Qy      1081 ATGTGCGGCGCAACGCGGCTCTTCGCAACCTTTCGCAACGCGGCGCAACGCGGCGCA 1140
Db      1099 ATGTGCGGCGCAACGCGGCTCTTCGCAACCTTTCGCAACGCGGCGCAACGCGGCGCA 1158
Qy      1141 AGAGAGGCGGCTCACTTATCTCTGCGGAGCCCACTGCAACAGAGGCGGCTTTCGAG 1200
Db      1159 AGAGAGGCGGCTCACTTATCTCTGCGGAGCCCACTGCAACAGAGGCGGCTTTCGAG 1218
Qy      1201 ACTTAAATGTATACCACTTAACTGTGAGAGGGGAGCCCAATCTGGAATCTTCCCGGCT 1260
Db      1219 ACTTAAATGTATACCACTTAACTGTGAGAGGGGAGCCCAATCTGGAATCTTCCCGGCT 1278
Qy      1261 TGGGACATTCGAGGCGGAGAGAGAGTGCAGGCTGCGGCGGAGAGAGTGCAGAA 1320
Db      1279 TGGGACATTCGAGGCGGAGAGAGAGTGCAGGCTGCGGCGGAGAGAGTGCAGAA 1338
Qy      1321 GGGGCACTGAGGCGGCTGTGCGGCGAGGCTTGGGACATTCGAGGAGCAAGGAAAGTCTCC 1380
Db      1339 GGGGCACTGAGGCGGCTGTGCGGCGAGGCTTGGGACATTCGAGGAGCAAGGAAAGTCTCC 1398
Qy      1381 TGGGCGCATCTGTAAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1421
Db      1399 TGGGCGCATCTGTAAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1439

```

```

RESULT 5
US-10-187-657-4
; Sequence 4, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Laeak, Amy K. W.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program

```

```

; SEQ ID NO 4
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 322161CA2
US-10-187-657-4

Query Match      99.9%; Score 1415.8; DB 5; Length 1420;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1417; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GAAACGAAAGTAAAGGCTTACAGACAGTGAAGAAATAGTTTCCTGCGCGGCTAGAAAAA 60
Db      1  GAAACGAAAGTAAAGGCTTACAGACAGTGAAGAAATAGTTTCCTGCGCGGCTAGAAAAA 60
Qy      61  CTCTGTGCTAACCAACCCGAGGCTTGAAGAGAGCCCACTTCCAGCTTCTTAACGGA 120
Db      61  CTCTGTGCTAACCAACCCGAGGCTTGAAGAGAGCCCACTTCCAGCTTCTTAACGGA 120
Qy      121  GAGTGCAGAGACTCAGACTTACCAAGCCCACTGGTCCAGGCTTGAAGAGAGAG 180
Db      121  GAGTGCAGAGACTCAGACTTACCAAGCCCACTGGTCCAGGCTTGAAGAGAGAG 180
Qy      181  CCAAGAGCGGCTCTCCGCGCTCAAGGAGCCCAAGCTTGTGAGCTTGTGCGCGCTG 240
Db      181  CCAAGAGCGGCTCTCCGCGCTCAAGGAGCCCAAGCTTGTGAGCTTGTGCGCGCTG 240
Qy      241  CGTGCAGACTCGGCGGCGGTGAGAGATGACCTGTGGAACGCGGTATCTGCTTTTACC 300
Db      241  CGTGCAGACTCGGCGGCGGTGAGAGATGACCTGTGGAACGCGGTATCTGCTTTTACC 300
Qy      301  CCGAGCCCGGCGATGCGCGAGGCTTCAAGGCTTCACTGCTCATCTGATCTAGTGT 360
Db      301  CCGAGCCCGGCGATGCGCGAGGCTTCAAGGCTTCACTGCTCATCTGATCTAGTGT 360
Qy      361  TGGCTCAGAGAGAGGCTTCTGCTCATCTTTCGCGGAGATCCGTGCGCACTGCGCTGT 420
Db      361  TGGCTCAGAGAGAGGCTTCTGCTCATCTTTCGCGGAGATCCGTGCGCACTGCGCTGT 420
Qy      421  TTTGTTGTGAGAGTTCTTCTCAGTCTGTTCATTAAGCGAGAAATGTGAGTGTGCACT 480
Db      421  TTTGTTGTGAGAGTTCTTCTCAGTCTGTTCATTAAGCGAGAAATGTGAGTGTGCACT 480
Qy      481  TCAGTGAGAAATGTTCTGTGAGTACGTGAACCAACATCTTAAGAGCTTTCAGCG 540
Db      481  TCAGTGAGAAATGTTCTGTGAGTACGTGAACCAACATCTTAAGAGCTTTCAGCG 540
Qy      541  CAGGCGGCTTACAGCCCGTGTGCTGCTGCTGAGGCTGAGAGGCAATTAATTAAC 600
Db      541  CAGGCGGCTTACAGCCCGTGTGCTGCTGCTGAGGCTGAGAGGCAATTAATTAAC 600
Qy      601  TCAAGGAGCCCAAGTGCATCAGCTGAACGAGACATTAACAGAGAGATTCACCT 660
Db      601  TCAAGGAGCCCAAGTGCATCAGCTGAACGAGACATTAACAGAGAGATTCACCT 660
Qy      661  GGGCTCTGAAAGAAATTACGCGCGGAGTACGGAACGCACTGGAAGAAAGGGGCTGCGG 720
Db      661  GGGCTCTGAAAGAAATTACGCGCGGAGTACGGAACGCACTGGAAGAAAGGGGCTGCGG 720
Qy      721  ACCCAGTGTCTACTCTGCGGAGAAAGTTCAACACCGAGTACCCCTTGGCGCTGTACCA 780
Db      721  ACCCAGTGTCTACTCTGCGGAGAAAGTTCAACACCGAGTACCCCTTGGCGCTGTACCA 780
Qy      781  AGTACCACTGCGGCGGAGCACTACGCTCGGCGACGCTATGAGTGGCTTCTGCTTGGC 840
Db      781  AGTACCACTGCGGCGGAGCACTACGCTCGGCGACGCTATGAGTGGCTTCTGCTTGGC 840
Qy      841  TCCCTCCAAAGTGTCTCTCCACGCGCGCGCTCTAAGAGGCTGTGCACTGTGTA 900
Db      841  TCCCTCCAAAGTGTCTCTCCACGCGCGCGCTCTAAGAGGCTGTGCACTGTGTA 900

```


Db 1003 GCTCGGCTAGAGCTCTCGGCGCTCACTGATAGGCGCCCTTCTGAGTCAAGCT 1062
Qy 1028 GGGACACCGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1087
Db 1063 GGGACACCGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1122
Qy 1088 GGGACACCGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1147
Db 1123 GGGACACCGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1182
Qy 1148 GGGCTCACT 1207
Db 1183 GGGCTCACT 1242
Qy 1208 ATGTATCACT 1267
Db 1243 ATGTATCACT 1302
Qy 1268 TGGCAGGCGCGGGAAGAGTCCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1327
Db 1303 TGGCAGGCGCGGGAAGAGTCCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1362
Qy 1328 GAGCGCTGCTGCGCGCGAGGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1387
Db 1363 GAGCGCTGCTGCGCGCGAGGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1422
Qy 1388 ATCTGTAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1421
Db 1423 ATCTGTAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1456

RESULT 7

US-10-187-657-9
; Sequence 9, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Laeak, Amy K. W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incycle ID No. US20030068311A1 110769_Mm.1
; US-10-187-657-9

Query Match 48.8%; Score 693; DB 5; Length 1594;
Best Local Similarity 75.9%; Pred. No. 4,2e-204;
Matches 914; Conservative 0; Mismatches 275; Indels 16; Gaps 4;

Qy 230 CTGCGCGCTCTGTCAGACATCGGCGCGCTGTCAGACATCGGCGCGCTGTCAGACAT 289
Db 107 CTGCGCGCTCTGTCAGACATCGGCGCGCTGTCAGACATCGGCGCGCTGTCAGACAT 166
Qy 290 GCGCTTTTACCCGAGCGCGCGAGCGCGAGCGCTTACGCGCTTACGCGCTTACGCGCT 349
Db 167 ACCCTTTTACCCGAGCGCGAGCGCGAGCGCTTACGCGCTTACGCGCTTACGCGCT 226
Qy 350 TCTAGT 409
Db 227 CTTGT 286

Qy 410 CTGCGCGCTGT 469
Db 287 CTGCGCGCTGT 346
Qy 470 GCGCTGTGCACTTCAAGTCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 529
Db 347 GCGCTGTGCACTTCAAGTCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 406
Qy 530 AGCTTTACGCGAGCGCGGCTTACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 589
Db 407 AGCTTTACGCGAGCGCGGCTTACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 466
Qy 590 TAATTAATTAACCTCAAGGAGACCCAGTGCATCACTGTAAGAGACCAATTAATTAACCA 649
Db 467 TAATTAATTAACCTCAAGGAGACCCAGTGCATCACTGTAAGAGACCAATTAATTAACCA 526
Qy 650 GCAGTTCACTGCGCTGTGTGAAGAGATTAACCGCGGAGTACGCGAGCACTGAGAA 709
Db 527 GCAGTTCACTGCGCTGTGTGAAGAGATTAACCGCGGAGTACGCGAGCACTGAGAA 586
Qy 710 GGGGCTGCGGAGACCCAGTGTCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769
Db 587 GGGGCTGCGGAGACCCAGTGTCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 646
Qy 770 CTTTACCAACAGTACCACTGCGGAGCACTAGGCTGCGGAGCACTAGGCTGCGGAG 829
Db 647 CTTTACCAACAGTACCACTGCGGAGCACTAGGCTGCGGAGCACTAGGCTGCGGAG 706
Qy 830 CTGCTTCTGCTCTCTCTCAACGTGTGTCTCTCAACGCGGCGCGCTCTTACGAGGCT 889
Db 707 CTGCTTCTGCTCTCTCTCAACGTGTGTCTCTCAACGCGGCGCGCTCTTACGAGGCT 766
Qy 890 GGCACGTGTACCAACCGGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 949
Db 767 GGCACGTGTACCAACCGGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 826
Qy 950 TAGGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1009
Db 827 TAGGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 886
Qy 1010 GCGCTTCTGAGTCAACGTGTGCAACCGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 1069
Db 887 GCGCTTCTGAGTCAACGTGTGCAACCGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 946
Qy 1070 GAGTCTTCAAGT 1129
Db 947 GAGTCTTCAAGT 1006
Qy 1130 CT---GGAGCAAGAGAGAGGCGCTCACTCTTATCTGTGGGAGACCACTGACAGAA 1186
Db 1007 CTGTAGCAACAGGCTTAAGAGAACTCACTCTTCACTTCAACCAACCGGAGAGAA 1066
Qy 1187 GGCAGCTCTCCAGACTTAATATGTATCAACACTTACCTGTGAGAGGAGAGCACTTGA 1246
Db 1067 GTTGAAGAGTCCAGACTTA---TATTACCACTCTCTGTGAATAATTAAGTCTCGGA 1123
Qy 1247 CTCTTCCCGCGCTTGGAGATGCGAGGCGCGGAGAGTGTCCCGCAGGCTGTG-GGCCA 1305
Db 1124 TTCTTCACTCTCTTGGAGACCCATACCTGTAAGTGTGTGAAGCGGTGTCCAGGA 1183
Qy 1306 GGAAGCTTCAGAAAGGAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1365
Db 1184 GCATAGCTTATGTGTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1243
Qy 1366 CC-----AGGAAAGTCTCTGTGGGCGATCTGTAAATTAACCTTTTCTTTTGT 1416
Db 1244 CTTTGTCTTAAAGAAATGGAATTTCTCAGAGAACTGTAAATTAACCTTTTGT 1303
Qy 1417 TTTTCT 1421
Db 1304 TTTTCT 1308

```
RESULT 8
US-10-187-657-5
; Sequence 5, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661R6
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 455, 480, 483
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match
Best Local Similarity 32.3%; Score 459.6; DB 5; Length 522;
Matches 510; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

1 GAACCGAGAAAGTAAAGCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGCTAGAAAA 60
1 GAACCGAGAAAGTAAAGCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGCTAGAAAA 60
61 CTCTGTCGGTACCAACCCGAGAGCGTTGAGAGAGCCGCTCCAGGCTTCTTAACGGA 120
61 CTCTGTCGGTACCAACCCGAGAGCGTTGAGAGAGCCGCTCCAGGCTTCTTAACGGA 120
121 GAGGTGCAAGACTCAGACTTCAACCAAGCCACTGCTCCAGCCTTGTAGCAAAAGAGAG 180
121 GAGGTGCAAGACTCAGACTTCAACCAAGCCACTGCTCCAGCCTTGTAGCAAAAGAGAG 180
181 CCAAGGACGCGCTCTCCCGGCTCAGGCAAGCCCAAGCTTGCTGCTGCTGCGGCTG 240
181 CCAAGGACGCGCTCTCCCGGCTCAGGCAAGCCCAAGCTTGCTGCTGCTGCGGCTG 240
241 CGTGGAGCACTGCGCGCGGCTGAGCATACCTGTGGAACGGCGGTAAGCTTTTACC 300
241 CGTGGAGCACTGCGCGCGGCTGAGCATACCTGTGGAACGGCGGTAAGCTTTTACC 300
301 CCCAGCCCCGAGCTGCGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 360
301 CCCAGCCCCGAGCTGCGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 360
361 TGGCTTGAAGAGAGCTTCTGCTCATCTTTCGCGGAGAACCTGAGCACTGCGGCTG 420
361 TGG-TCAGAGAGAG--TTCGCTCATCTTTCGCGGAGAACCTGAGCACTGCGGCTG 417
421 TTTGTTGGTGAAGAGTTCTTCTGAGCTGTCTTCAATAGGCGAGAAATTTGGCTGCACT 480
421 TTTGTTGGTGAAGAGTTCTTCTGAGCTGTCTTCAATAGGCGAGAAATTTGGCTGCACT 477
481 TCAGTCAGAAATGGTCTGCTGAGTACAGTGAACCAACATCTTAC 527
481 TCAGTCAGAAATGGTCTGCTGAGTACAGTGAACCAACATCTTAC 527
478 TCGGT-NAGATGGTTCTGCTGAGTACAGTGAACCAACATCTTAC 522
```

```
RESULT 9
US-10-773-236-69
; Sequence 69, Application US/10773236
; Publication No. US20050208602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et. al.
; TITLE OF INVENTION: 89 Human Secreted Proteins
; FILE REFERENCE: P5751P1
; CURRENT APPLICATION NUMBER: US/10/773,236
; PRIOR FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: 60/311,085
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,209
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/US02/25107
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/330,629
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US02/33985
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/331,046
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: PCT/US02/35606
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/358,554
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: PCT/US03/04819
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/358,714
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 396
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 69
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-773-236-69

Query Match
Best Local Similarity 29.1%; Score 413; DB 9; Length 580;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAACCGAGAAAGTAAAGCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGCTAGAAAA 60
5 GAACCGAGAAAGTAAAGCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGGCTAGAAAA 64
61 CTCTGTCGGTACCAACCCGAGAGCGTTGAGAGAGCCGCTCCAGGCTTCTTAACGGA 120
65 CTCTGTCGGTACCAACCCGAGAGCGTTGAGAGAGCCGCTCCAGGCTTCTTAACGGA 124
121 GAGGTGCAAGACTCAGACTTCAACCAAGCCACTGCTCCAGCCTTGTAGCAAAAGAGAG 180
125 GAGGTGCAAGACTCAGACTTCAACCAAGCCACTGCTCCAGCCTTGTAGCAAAAGAGAG 184
181 CCAAGGACGCGCTCTCCCGGCTCAGGCAAGCCCAAGCTTGCTGCTGCTGCGGCTG 240
185 CCAAGGACGCGCTCTCCCGGCTCAGGCAAGCCCAAGCTTGCTGCTGCTGCGGCTG 244
241 CGTGGAGCACTGCGCGCGGCTGAGCATACCTGTGGAACGGCGGTAAGCTTTTACC 300
245 CGTGGAGCACTGCGCGCGGCTGAGCATACCTGTGGAACGGCGGTAAGCTTTTACC 304
301 CCCAGCCCCGAGCTGCGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 360
305 CCCAGCCCCGAGCTGCGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 364
361 TGGCTTGAAGAGAGCTTCTGCTCATCTTTCGCGGAGAACCTGAGCACTGCGGCTG 413
365 TGGCTTGAAGAGAGCTTCTGCTCATCTTTCGCGGAGAACCTGAGCACTG 417
```

RESULT 10

US-10-450-763-2367/c
; Sequence 2367, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 2367
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (229)..(92)
; OTHER INFORMATION: 41% homologous to Homo sapiens Human secreted protein, SEQ ID
; OTHER INFORMATION: NO: 7589, accession number G03508, Smith-Waterman Score=71.
US-10-450-763-2367

Query Match 28.3%; Score 402.8; DB 9; Length 406;
Best Local Similarity 99.5%; Pred. No. 3,4e-114; Indels 0; Gaps 0;
Matches 404; Conservative 0; Mismatches 2;

QY 8 GAAAGTACCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGCTAGAAAACCTGTGC 67
DB 406 GAAAGTACCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGCTAGAAAACCTGTGC 347
QY 68 GGTACCAACCCAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAAGGAGAGTGC 127
DB 346 GGGGCAACCCAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAAGGAGAGTGC 287
QY 128 AGGAGTACAGCTTACAGAGCCCACTGCTCCAGCTTGTAGCAAGAGAGAGAGAGGA 187
DB 286 AGGAGTACAGCTTACAGAGCCCACTGCTCCAGCTTGTAGCAAGAGAGAGAGAGGA 227
QY 188 CGGCTCTCCCGCTGCGAGGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 247
DB 226 CGGCTCTCCCGCTGCGAGGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 167
QY 248 CACTCGCGCGCGCTGCGAGCATGACCTGTGGAAGCGGCTACTGCTTTTAAACCCAGCC 307
DB 166 CACTCGCGCGCGCTGCGAGCATGACCTGTGGAAGCGGCTACTGCTTTTAAACCCAGCC 107
QY 308 CCGGCAAGCGCGAGGCTTACAGCGTTCACATGCTCATGTTATTAAGTGTGTTGGCTCT 367
DB 106 CCGGCAAGCGCGAGGCTTACAGCGTTCACATGCTCATGTTATTAAGTGTGTTGGCTCT 47
QY 368 AGCAGCAAGCTTCTGCTCATCTTGCAGGAGATCCGTCGACCACTCG 413
DB 46 AGCAGCAAGCTTCTGCTCATCTTGCAGGAGATCCGTCGACCACTCG 1

RESULT 11
US-09-925-065A-72673/c
; Sequence 72673, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096

;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 72673
;; LENGTH: 1376
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-72673

Query Match 27.9%; Score 396.6; DB 4; Length 1376;
Best Local Similarity 99.7%; Pred. No. 3.9e-112;
Matches 396; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACCGAAGAAAGTACCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGCTAGAAA 60
DB 397 GAACCGAAGAAAGTACCGCTACAGACAGTGAAGAAATAGTTGCTGCGCGCTAGAAA 338
QY 61 CTCTGTGCTTACCAACCCAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAACGGA 120
DB 337 CTCTGTGCTTACCAACCCAGAGCGTTGAGAGAGCCCACTTCCAGCTTCTTAACGGA 278
QY 121 GAGTGCAGAGACTCAGACTTACAGAGCCCACTGCTCCAGCTTGTAGCAAGAGAGAG 180
DB 277 GAGTGCAGAGACTCAGACTTACAGAGCCCACTGCTCCAGCTTGTAGCAAGAGAGAG 218
QY 181 CCAAGAGCGCTCTCTCCCGCTGCGAGGAGCCCACTTGTGCTGCTGCTGCTGCTG 240
DB 217 YCAAGAGCGCTCTCTCCCGCTGCGAGGAGCCCACTTGTGCTGCTGCTGCTGCTG 158
QY 241 CGTGCAAGACTCGGCGCGCGCTGAGAGATGACCCGTGGAAGCGGCTACTGCTTTTAA 300
DB 157 CGTGCAAGACTCGGCGCGCGCTGAGAGATGACCCGTGGAAGCGGCTACTGCTTTTAA 98
QY 301 CCAAGCGCGGAGTGCAGAGGCTTACAGGCTTCACTGCTCATGTTATTAAGTGTGTT 360
DB 97 CCAAGCGCGGAGTGCAGAGGCTTACAGGCTTCACTGCTCATGTTATTAAGTGTGTT 38
QY 361 TGGCTTAGAGCAAGCTTCTGCTCATCTTGCAGG 397
DB 37 TGGCTTAGAGCAAGCTTCTGCTCATCTTGCAGG 1

RESULT 12
US-10-094-749-1195/c
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHITO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYU
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1195
LENGTH: 2684
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-1195

Query Match 27.5%; Score 391; DB 6; Length 2684;
Best Local Similarity 97.5%; Pred. No. 2,5e-110;
Matches 397; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1015 TCTGGGTACGCTGGCAACCGGCTCTGCTCTTCCCGGAGGGGCGCTGAGATC 1074
DB 2363 TTTCATCCCGACCGACAGGCGTCTGCTCTTCTCGAGGGGCGGTGTAATC 2304
QY 1075 TCCAGTATGTTCCGCCAGCGCTCTTCCGACCCCTTCGACCAAGCGCCAGACTGCA 1134
DB 2303 TCCAGTATGTTCCGCCAGCGCTCTTCCGACCCCTTCGACCAAGCGCCAGACTGCA 2244
QY 1135 GCCAGAGAGAGGGGCTCACTCTTATCTCGGCCAACCACCTGCAAGAGAGCGCTC 1194
DB 2243 GCCAGAGAGAGGGGCTCACTCTTATCTCGGCCAACCACCTGCAAGAGAGCGCTC 2184
QY 1195 TCCCACTTAAATGATACACCACTAACCTGAGGGGGAGCCCAATCTGACTCTCTCC 1254
DB 2183 TCCCACTTAAATGATACACCACTAACCTGAGGGGGAGCCCAATCTGACTCTCTCC 2124
QY 1255 CCGCTTGGGACATCGCAAGCGCCGGAAGAGAGTCCCGCCAGGCTGAGGAGAGCTC 1314
DB 2123 CCGCTTGGGACATCGCAAGCGCCGGAAGAGAGTCCCGCCAGGCTGAGGAGAGCTC 2064
QY 1315 CAGGAAGGGGACATGAGCGCTGCTGGGCGGAGGCTTGGAGATCCGAGGCAAGGAGAA 1374
DB 2063 CAGGAAGGGGACATGAGCGCTGCTGGGCGGAGGCTTGGAGATCCGAGGCAAGGAGAA 2004
QY 1375 GTCTCTGGGGCGATCTGTAATAAATCTTTTCTTTTCTTTT 1421
DB 2003 GTCTCTGGGGCGATCTGTAATAAATCTTTTCTTTTCTTTT 1957

RESULT 13
US-10-187-657-7
Sequence 7, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT APPLICATION NUMBER: US/10/187,657
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: PCT/US00/07817
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20030068311A1 SEQID4510D1
FEATURE:
NAME/KEY: unsure
LOCATION: 140, 165, 203, 237, 269, 289, 307, 338, 344, 347, 354, 390, 430, 445
OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-7

Query Match 27.1%; Score 385.8; DB 5; Length 450;
Best Local Similarity 93.7%; Pred. No. 6.6e-109;
Matches 404; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 624 CTGAACGAGACCATGACTACACAGAGCAGTTCACCTGCGCTGTGAAGAATTAAGCC 683
DB 21 CTGAACGAGACCATGACTACACAGAGCAGTTCACCTGCGCTGTGAAGAATTAAGCC 80
QY 684 GCGAGTACGCGACGACCTGAGAGAGGGGCTGCCGAAACCAAGTGTCTTACTGCGGAG 743
DB 81 GCGAGTACGCGACGACCTGAGAGAGGGGCTGCCGAAACCAAGTGTCTTACTGCGGAG 140
QY 744 AAGTTCACACGAGTACAGCCCTTGCGGCTGTACACACAGTACCACTGCGGGAGACATAC 803
DB 141 AAGTTCACACGAGTACAGCCCTTGCGGCTGTACACACAGTACCACTGCGGGAGACATAC 200
QY 804 GCTCGGCCACGCTATGAGTGGCGTTCCTTCTGCTCTCTCCAAAGTGTCTCTCC 863
DB 201 GNTGCGGCCACGCTATGAGTGGCGTTCCTTCTGCTCTCTCCAAAGTGTCTCTCC 260
QY 864 AGCGCGGCGCGCTCTACGAGAGGCTGAGACCTGTGACCAACGAGAGCTTCCGCTCTTC 923
DB 261 AGCGCGGCGCGCTCTACGAGAGGCTGAGACCTGTGACCAACGAGAGCTTCCGCTCTTC 320
QY 924 GGGGTCTTGGCTTGGCTCTTCCATCTTACGAGGCTGTGCGCGCTCGGCTCGGCTCTCC 983
DB 321 GGGGTCTTGGCTTGGCTCTTCCATCTTACGAGGCTGTGCGCGCTCGGCTCGGCTCTCC 380
QY 984 TCCGCGCTGACCACTGAGTACGAGCGCGCTTCTGAGTACGCTGTGCAACGAGCTCTG 1043
DB 381 TCCGCG-TCANCAATCAGTACGAGACGAGCTTGTGAGTCAAGCTGTGCAANAGCGCTG 439
QY 1044 TGCCTTCTCT 1054
DB 440 TGCCTTCTCT 450

RESULT 14
US-09-759-130B-424
Sequence 424, Application US/09759130B
Publication No. US20030022279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A.
APPLICANT: Fraser, Christopher C.
APPLICANT: Sharp, John D.
APPLICANT: Barnes, Thomas S.
APPLICANT: Kirtel, Susan J.
APPLICANT: Mackay, Charles R.
APPLICANT: Meyers, Paul S.
APPLICANT: Leiby, Kevin R.
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodheart, Andrew
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
FILE REFERENCE: MP100-5350NM1M
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24

PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/596,194
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 09/342,364
 PRIOR FILING DATE: 1999-06-29
 PRIOR APPLICATION NUMBER: US 09/608,452
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/393,996
 PRIOR FILING DATE: 1999-09-10
 PRIOR APPLICATION NUMBER: US 09/602,871
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 09/420,707
 PRIOR FILING DATE: 1999-10-19
 NUMBER OF SEQ ID NOS: 460
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 424
 LENGTH: 1029
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-759-130B-424

Query Match 25.0%; Score 354.8; DB 3; Length 1029;
 Best Local Similarity 64.5%; Pred. No. 3,4e-99;
 Matches 546; Conservative 0; Mismatches 297; Indels 3; Gaps 1;

287 ACTGCTTTTACCCCGAGCCCGGCAATGCCGAGCTTCAAGCTTCACTGCTCATCGT 346
 21 ATTCCCTTCTATGCTGCGCCCAAGCCCACTTCCCATGACACCACTTTGGCGACAT 80
 347 TATTCTAGTGTGCTGCTAGACAGAGCTTCTGCTCATCTGCGGGGATCCGTG 406
 81 CATCATGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
 407 CCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
 141 AAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 200
 467 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
 201 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 260
 527 CAAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 586
 261 CAGGCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
 587 CATTAATATTACTCACTCAAGGAGCCCGAGTGCATCACTGACCAAGCAATTGACTCAA 646
 321 AGTCAATCACTCACTCAAGGAGCCCGAGTGCATCACTGACCAAGCAATTGACTCAA 380
 647 CGAGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
 381 CGAGGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
 707 GAAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
 441 GAAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
 767 CGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826
 501 TGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
 827 GTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
 561 ATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 887 CCGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
 621 CTACATGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
 944 CATCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
 681 CACATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 740

1004 CGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
 741 TGGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
 1064 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
 801 TATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
 1124 CAGGA 1129
 861 GAGTGA 866

RESULT 15
 US-10-042-431-54
 Sequence 54, Application US/10042431
 Publication No. US20020182675A1

GENERAL INFORMATION:
 APPLICANT: MCCARTHY, Sean A
 APPLICANT: BARNES, Thomas M
 APPLICANT: PRASER, Christopher C
 APPLICANT: SHARP, John D
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 PREVENTIVE, THERAPEUTIC, AND OTHER USES
 FILE REFERENCE: 10147-602
 CURRENT APPLICATION NUMBER: US/10/042,431
 CURRENT FILING DATE: 2001-10-25
 PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/578,063
 PRIOR FILING DATE: 2000-05-24
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 54
 LENGTH: 1029
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-042-431-54

Query Match 25.0%; Score 354.8; DB 5; Length 1029;
 Best Local Similarity 64.5%; Pred. No. 3,4e-99;
 Matches 546; Conservative 0; Mismatches 297; Indels 3; Gaps 1;

287 ACTGCTTTTACCCCGAGCCCGGCAATGCCGAGCTTCAAGCTTCACTGCTCATCGT 346
 21 ATTCCCTTCTATGCTGCGCCCAAGCCCACTTCCCATGACACCACTTTGGCGACAT 80
 347 TATTCTAGTGTGCTGCTAGACAGAGCTTCTGCTCATCTGCGGGGATCCGTG 406
 81 CATCATGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
 407 CCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
 141 AAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 200
 467 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
 201 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 260
 527 CAAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 586
 261 CAGGCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
 587 CATTAATATTACTCACTCAAGGAGCCCGAGTGCATCACTGACCAAGCAATTGACTCAA 646
 321 AGTCAATCACTCACTCAAGGAGCCCGAGTGCATCACTGACCAAGCAATTGACTCAA 380
 647 CGAGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
 381 CGAGGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
 707 GAAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766

```
Db      441 GAAGGGGCTGGCAGACCTGTGTGTGTAAGCTGAGAAAGTTCACTCCAAAGAACCCATG 500
QY      767 CGGCGCTGTACCAACAGATACCACTGGCGGGAGACATAAGCTTCGGCCACGCTATGGGTGGC 826
Db      501 TGGCCCTATACCCGCAAGTACCGGCTGGCGGGACATACTCAGCCATGCTATGGGTGGC 560
QY      827 GTTCTGCTTCTGAGCTCCCTCTCCAAAGTGTGCTCTCCAGCCGGCCCGGCTCTAAGGAAG 886
Db      561 ATTCTCTGCTGAGCTGGCTGGCCATGTGATGCTCTCCATGCTGTGCTGTGTATATGTGG 620
QY      887 CCTGGCACTGTGACCAACCGAGCCTTTCGCGCTTTG--GAGTCTTGCCCTTGGCTTC 943
Db      621 CTACATGCTATTTGGCCACGGGCAATCTTCCAGCTGTGGCTGTGCTCTTCTTCCATGGC 680
QY      944 CATCTTGAAGGTGCCGCTTGGCCGCTCCGCTAAGGCTCTCCGCGCTCACCACTAGTA 1003
Db      681 CACATCACTCACCTCAACCTGTCCCTGTGCACTGGGCGCTTGTGTGCTGATATCAACCA 740
QY      1004 CGGCGCGGCTTCTGAGTCAAGCTGGCAACCGGCGTCTGTGCTCTTCTCGGAGGGGC 1063
Db      741 TGGGCTGTGCTTGTGATCACATTGACCAAGGACTGTGTGTGCTGTGGCTTGGC 800
QY      1064 CGTGTGAGTCTTCAGATATGTTCGAGCCAGCGCTTTCGACCCCTTCTGAGCCAAAGCGC 1123
Db      801 TATGGCGGTGGCCCAAGGATGAGGCTCACAGGCTGAAGGCTTTCTTCAACCAAGAGTGT 860
QY      1124 CAAGCA 1129
Db      861 GGATGA 866
```

Search completed: January 22, 2006, 03:01:37
Job time : 1288 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 22:44:31 ; Search time 320 Seconds
(without alignments)
3660.288 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccgggaagaacgaagct.....ctttttctttctttttt 1421

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6049916 seqs, 412136615 residues

Total number of hits satisfying chosen parameters: 12099832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA New:
1: /cgn2_6/ptodaca/1/pubpna/US06_NEW_PUB.seq:*
2: /cgn2_6/ptodaca/1/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodaca/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodaca/1/pubpna/US09_NEW_PUB.seq:*
5: /cgn2_6/ptodaca/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodaca/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodaca/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodaca/1/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodaca/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodaca/1/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/ptodaca/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.8	3.2	2319	US-11-136-527-1934	Sequence 1934, App
2	43	3.0	1006	US-10-821-234-422	Sequence 422, App
3	43	3.0	1071	US-11-000-688-1164	Sequence 1164, App
4	42.6	3.0	88421	US-11-205-109-1	Sequence 1, Appli
5	41.8	2.9	88421	US-11-205-109-1	Sequence 1, Appli
6	41.6	2.9	1400	US-11-136-527-4440	Sequence 4440, App
7	41.6	2.9	2305	US-11-136-527-344	Sequence 344, App
8	41.6	2.9	12309	US-10-995-561-13450	Sequence 13450, A
9	41.6	2.9	95832	US-10-995-561-13273	Sequence 13273, A
10	41.4	2.9	173602	US-11-121-086-25	Sequence 25, Appli
11	41.4	2.9	2432	US-10-750-185-29908	Sequence 29908, A
12	41.4	2.9	2432	US-10-750-185-29908	Sequence 29908, A
13	41.4	2.9	3886	US-11-087-100-37	Sequence 37, Appli
14	41.4	2.9	3886	US-11-087-100-37	Sequence 37, Appli
15	41.4	2.9	3886	US-11-087-100-37	Sequence 37, Appli
16	40.8	2.9	2862	US-11-037-243-7	Sequence 12765, A
17	40.4	2.8	201	US-10-995-561-12765	Sequence 224, App
18	40.4	2.8	11438	US-10-821-234-422	Sequence 17, Appli
19	40	2.8	2797	US-11-075-400-17	Sequence 499, App
20	40	2.8	11612	US-10-995-561-499	Sequence 113, App
21	39.8	2.8	717	US-10-883-512-113	Sequence 109, App
22	39.8	2.8	895	US-10-883-512-109	Sequence 109, App

23	39.8	2.8	895	US-10-883-512-112	Sequence 112, App
24	39.8	2.8	955	US-10-883-512-110	Sequence 110, App
25	39.8	2.8	34875	US-10-775-169-316	Sequence 316, App
26	39.8	2.8	164810	US-11-121-086-4	Sequence 4, Appli
27	39.8	2.8	172543	US-11-121-086-6	Sequence 6, Appli
28	39.8	2.8	2133	US-11-143-980-13	Sequence 13, Appli
29	39.2	2.8	116856	US-11-143-980-13	Sequence 1, Appli
30	38.8	2.7	10800	US-11-019-711-5	Sequence 5, Appli
31	38.8	2.7	10809	US-11-019-711-1	Sequence 1, Appli
32	38.4	2.7	1494	US-11-143-980-26	Sequence 26, Appli
33	38.4	2.7	157224	US-11-112-908-51	Sequence 51, Appli
34	38.4	2.7	161726	US-11-112-908-48	Sequence 48, Appli
35	38.4	2.7	161726	US-11-112-908-52	Sequence 52, Appli
36	38.4	2.7	170189	US-11-112-908-50	Sequence 50, Appli
37	38.4	2.7	1850	US-10-883-512-114	Sequence 114, App
38	38.2	2.7	3200	US-10-645-441-13	Sequence 13, Appli
39	38	2.7	1110	US-10-432-483-13	Sequence 14, Appli
40	38	2.7	1897	US-10-775-169-236	Sequence 236, App
41	38	2.7	8651	US-10-432-483-48	Sequence 48, Appli
42	37.8	2.7	963	US-10-858-730-182	Sequence 182, App
43	37.8	2.7	1725	US-10-821-234-772	Sequence 772, App
44	37.8	2.7	1883	US-10-821-234-772	Sequence 9, Appli
45	37.2	2.6	1527	US-11-073-185-56	Sequence 56, Appli

ALIGNMENTS

RESULT 1
US-11-136-527-1934
; Sequence 1934, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1934
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-1934

Query Match 3.2%; Score 44.8; DB 8; Length 2319;
Best Local Similarity 51.5%; Pred. No. 0.085;
Matches 103; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY GCTCTCCACCGCCGCGCTCTACGAGGCTTGCACTGCTGACCAACGAGCTTGGC 916
DB 776 GCAACCGCTGTGTCTGTGCGACGCGGCTGTCTGTCTGTGCGGAGACT 835
QY 917 GCTCTCGAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 976
DB 836 GCGATCTCCACATCTGCAAGGCTTCATCCGCGCCAGCTCATCTGCGCCGAGAC 895
QY 977 AGGCTCTCGCGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 1036
DB 896 GCTGAGAGGAGAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 955
QY 1037 GCTCTGAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1056
DB 956 GAGACGACGCT 975
RESULT 2
US-10-821-234-422
; Sequence 422, Application US/10821234

Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmati, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: pt_seq_genes Version 1.0
SEQ ID NO 422
LENGTH: 1006
TYPE: DNA
ORGANISM: Homo sapiens
US-10-821-234-422

Query Match 3.0%; Score 43; DB 7; Length 1006;
Best Local Similarity 48.9%; Pred. No. 0.18;
Matches 115; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 722 CCCAGTCTCTACCTGGCGGAGAGTTTCACACCGAGTAGCCCTTGCGGCTGTACACCA 781
DB 295 CCGTGTCTACAGCTGTCCGAGAGGCGACACTGGGCCCTGTGTGCGCCCTGTGCTG 354
QY 782 GTACCACTGTGGCGGACACTACGCTTGGCCAGCTATGGGTGGCGTTGCTTGGCT 841
DB 355 GCAGCCGTTGACCGCGACGTTGACCGGGAACCTCTGCGACGTGGCGGCTGGGGCAT 414
QY 842 CCTCTCAAGTGTCTCTCTCCACGCGCGCCGCTTACGAGGCTGGCACTGTGAC 901
DB 415 AGTCAACACGCGCGCGCGCCCGGACAGCTTGCAGACGTCTTGCAGTGTGGA 474
QY 902 CACCGAGCCTTGGCGCTCTTGGGGGTCTTGGCCTTGCCTCATCTAGCGTG 956
DB 475 CCGCGCACCTGTGAACCGCGGACGACACGACGCGCGCATATCCGAGGCTTG 529

RESULT 3
US-11-000-688-1164
Sequence 1164, Application US/11000688
Publication No. US20050287544A1
GENERAL INFORMATION:
APPLICANT: BERTUCCI, Francois
APPLICANT: HOULGATTE, Rami
APPLICANT: BIRNBAUM, Daniel
TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
FILE REFERENCE: 1423-R-03
CURRENT APPLICATION NUMBER: US/11/000,688
CURRENT FILING DATE: 2004-12-01
PRIOR APPLICATION NUMBER: US 60/525,987
PRIOR FILING DATE: 2003-12-01
NUMBER OF SEQ ID NOS: 1596
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1164
LENGTH: 1071
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial sequences:primer
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1071)
OTHER INFORMATION: d component of complement (adipsin) (DF)
OTHER INFORMATION: gene.
US-11-000-688-1164

Query Match 3.0%; Score 43; DB 8; Length 1071;
Best Local Similarity 48.9%; Pred. No. 0.18;
Matches 115; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 722 CCCAGTCTCTACCTGGCGGAGAGTTTCACACCGAGTAGCCCTTGCGGCTGTACACCA 781
DB 324 CCGTGTCTACAGCTGTCCGAGAGGCGACACTGGGCCCTGTGTGCGCCCTGTGCTG 383
QY 782 GTACCACTGTGGCGGACACTACGCTTGGCCAGCTATGGGTGGCGTTTGTCTTGTGCT 841
DB 384 GCAGCCGTTGACCGCGACGTTGACCGGAACTCTGTGCACTGTGGCGGCTGGGGCAT 443
QY 842 CCTCTCAAGTGTCTCTCTCCACGCGCGCCGCTTACGAGGCTGGCACTGTGAC 901
DB 444 AGTCAACACGCGCGCGCGCCCGGACAGCTTGCAGACGTCTTGCAGTGTGGA 503
QY 902 CACCGAGCCTTGGCGCTCTTGGGGGTCTTGGCCTTGCCTCATCTAGCGTG 956
DB 504 CCGCGCACCTGTGAACCGCGGACGACACGACGCGCGCATATCCGAGGCTTG 558

RESULT 4
US-11-205-109-1/C
Sequence 1, Application US/11205109
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANTIN BIOSYNTHESIS
FILE REFERENCE: 3002-2US
CURRENT APPLICATION NUMBER: US/11/205,109
CURRENT FILING DATE: 2005-08-17
PRIOR APPLICATION NUMBER: US/09/976,059
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/239,924
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 88421
TYPE: DNA
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2077)..(3078)
OTHER INFORMATION: ORF 1; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3118)..(4032)
OTHER INFORMATION: ORF 2; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (6665)..(5814)
OTHER INFORMATION: ORF 4; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (7703)..(6693)
OTHER INFORMATION: ORF 5; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9464)..(8130)
OTHER INFORMATION: ORF 6; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9691)..(10761)
OTHER INFORMATION: ORF 7; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (12751)..(10829)
OTHER INFORMATION: ORF 8; negative strandedness
FEATURE:

```

NAME/KEY: misc feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 9; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 10; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15591)..(15863)
OTHER INFORMATION: ORF 11; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15880)..(119035)
OTHER INFORMATION: ORF 12; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (19032)..(39713)
OTHER INFORMATION: ORF 13; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 16; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70059)..(70662)
OTHER INFORMATION: ORF 18; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75424)..(74213)
OTHER INFORMATION: ORF 22; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc feature

```

```

LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (87494)..(88420)
OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only
US-11-205-109-1

```

```

Query Match 3.0%; Score 42.6; DB 8; Length 88421;
Best Local Similarity 46.5%; Pred. No. 1.4;
Matches 138; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

```

```

QY 839 GCTCTCTCCAGAGTGTCTGTCTCCAGCGCGCGCTCTAGAGAGCTGACATGCT 898
DB 13270 GGTGGCGCTGGCGGTCTTATGTGGACCTGGCGCGCGCTGACCGCGCGCGCA 13211
QY 899 GACCCAGAGAGCTTGGCGCTTCTCGGAGTCTTGGCTTGGCTTCTTACCTGAGTCC 958
DB 13210 GACCCAGAGAGCTTGGCGCTTCTCGGAGTCTTGGCTTGGCTTCTTACCTGAGTCC 13151
QY 959 GCTTGGCGCTTGGCGCTTCTCGGAGTCTTGGCTTGGCTTCTTACCTGAGTCC 1018
DB 13150 CGGACCGCGGTCTTGAAGATGTTCTCCCGGACGCTGACGACGACGCTGCGCG 13091
QY 1019 GGTTCAGCTTGGCGAGCGCGCTTCTCGGAGTCTTCTCGGAGGCGGTGAGTCC 1078
DB 13090 CGAGTGGCTTGAAGATGTTCTCCCGGACGCTGACGACGACGCTGCGCGCG 13031
QY 1079 GTATGTTGGCGAGCGCTTCTCGGAGTCTTCTCGGAGGCGGTGAGTCC 1135
DB 13030 GCGCGAGCTTACCGGTGACCGGCGCGCGCTTCTCGGAGTCTTCTCGGAGTCC 12974

```

```

RESULT 5
US-11-205-109-1
Sequence 1, Application US/11205109
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
TITLE OR INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
FILE REFERENCE: 3002-2US
CURRENT FILING DATE: US/11/205, 109
PRIOR APPLICATION NUMBER: 2005-08-17
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US/09/976, 059
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 88421

```

```

TYPE: DNA
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (2077)..(3078)
OTHER INFORMATION: ORF 1; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (3118)..(4032)
OTHER INFORMATION: ORF 2; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (6665)..(5814)
OTHER INFORMATION: ORF 4; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (7703)..(6693)
OTHER INFORMATION: ORF 5; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (9464)..(8130)
OTHER INFORMATION: ORF 6; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (9691)..(10761)
OTHER INFORMATION: ORF 7; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (12751)..(10829)
OTHER INFORMATION: ORF 8; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 9; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 10; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15591)..(13863)
OTHER INFORMATION: ORF 11; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15880)..(19035)
OTHER INFORMATION: ORF 12; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (19032)..(39713)
OTHER INFORMATION: ORF 13; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 16; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70099)..(70662)

```

```

OTHER INFORMATION: ORF 18; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75424)..(74213)
OTHER INFORMATION: ORF 22; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75335)..(76464)
OTHER INFORMATION: ORF 23; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (81674)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (87494)..(88420)
OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only
US-11-205-109-1

```

Query Match 2.9%; Score 41.8; DB 8; Length 88421;
 Best Local Similarity 50.8%; Pred. No. 2.2; 97; Indels 0; Gaps 0;
 Matches 100; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```

QY 805 CTTGGGACGATAGGAGCGTCTGCTTCTGCTCTCCACAGTGTGCTCTCA 864
DB 73052 CGTCAGAGCCACGCGGCGGCGTCTCCAGCAGATCTGTGAACCGGCGGCTCCA 73111
QY 865 CGCGGCGCGGCTCTACGAGGCGTGGCACTGTGACCAACGAGCGTTCGCGCTTCG 924

```

Db 73112 CCTGAAAGCGTGGGAGTCCGCGCCATCCGGGGCGAGGTGGGAAGTGAACGTCC 731171
Qy 925 GGGCTTCGCTTGGGCTTCATATCTAAGGTGCGGCTTGGCCGCTCCGCTAAGCTCCT 984
Db 73172 ACGGCTCGGACTGAGTGGCCCACTTGAAGTGGCCGCGCTTGGCCATGAACCCGCGCT 73231
Qy 985 CCGGCTCAACCACTGAG 1001
Db 73232 TGGCGATCTGATGCGG 73248

RESULT 6

US-11-136-527-4440/C
; Sequence 4440, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4440
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4440

Query Match 2.9%; Score 41.6; DB 8; Length 1400;
Best Local Similarity 35.3%; Pred. No. 0.47;
Matches 77; Conservative 42; Mismatches 99; Indels 0; Gaps 0;

Qy 845 CTCGAAGTGTCTGTCTCAAGCGCGGCGCTCTAGAGAGGCGCTGACAGTGAACAC 904
Db 796 CTRATCTKCKTKVKAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
Qy 905 CGAGGCTTCGCGCTTCTGCGGGCTTTCGCTTGGCTTCATCTAAGTGGCGCTCTG 964
Db 736 CTCTATATCTKKGATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
Qy 965 CCGGCTCGGCTTGAAGCTCTTCCGCGCTCAACCACTGAGTGAAGGCGCGCTTCTGAGTCA 1024
Db 676 CWRGCTCATATTTGDDAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
Qy 1025 GCTGGCAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
Db 616 TCTSKCT 579

RESULT 7

US-11-136-527-344/C
; Sequence 344, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 344
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-344

Query Match 2.9%; Score 41.6; DB 8; Length 2305;
Best Local Similarity 35.3%; Pred. No. 0.57;
Matches 77; Conservative 42; Mismatches 99; Indels 0; Gaps 0;

Qy 845 CTCGAAGTGTCTGTCTCAAGCGCGGCGCTCTAAGAGGCGCTGACAGTGAACAC 904
Db 1701 CTRATCTKCKTKVKAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642
Qy 905 CGAGGCTTCGCGCTTCTGCGGGCTTTCGCTTGGCTTCATCTAAGTGGCGCTCTG 964
Db 1641 CTCTATATCTKKGATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
Qy 965 CCGGCTCGGCTTGAAGCTCTTCCGCGCTCAACCACTGAGTGAAGGCGCGCTTCTGAGTCA 1024
Db 1581 CWRGCTCATATTTGDDAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522
Qy 1025 GCTGGCAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
Db 1521 TCTSKCT 1484

RESULT 8

US-10-995-561-13450
; Sequence 13450, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13450
; LENGTH: 12309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13450

Query Match 2.9%; Score 41.6; DB 7; Length 12309;
Best Local Similarity 49.1%; Pred. No. 1.1;
Matches 107; Conservative 1; Mismatches 110; Indels 0; Gaps 0;

Qy 907 GAGCTTCGCGCTTCTGCGGGCTTTCGCTTGGCTTCATCTAAGTGGCGCTCTGCC 966
Db 6026 GGGTCCCCGGGCGCTGCGAGGCTGAGTGGCGCCCGATGCGCGGCTTCCCGCGCGG 6085
Qy 967 CGCTCGGCTAAGCTCTTCCGCGCTCAACCACTGAGTGAAGGCGCGCTTCTGAGTCA 1026
Db 6086 CCGTCCCGGGAACGCGCGCGCTGCGCGCGGCTGCGCGCGCGCGCGCGCGCTGCG 6145
Qy 1027 TGGCAACCGGCTGCTGCTGCTTCTGCGAGGCGCGGCTGAGTGAAGTCAAGTATTC 1086
Db 6146 CGCTCATCTGAGCCCGCGCGCTGCGCGCGCGCGCGCGCGCGCTTCTAATGAGCCC 6205
Qy 1087 GAGCCAGCGCTTCTGCAACCTTCTGAGCAACCAAGCGCT 1124
Db 6206 CGGCGCGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6243

RESULT 9

US-10-995-561-13273/C
; Sequence 13273, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559

TYPE: DNA
ORGANISM: Bovine 1986680919140
US-10-750-623-29908

Query Match 2.9%; Score 41; DB 7; Length 2432;
Best Local Similarity 47.2%; Pred. No. 0.84;
Matches 125; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 788 CTTGGCGGAGCACTACGCTTGGGCAAGCTATGAGTGGCGCTTGTCTTGGCTCTCTC 847
DB 2073 CTTGGCGGAGGAGGAGGATGCGCGGCAAGCTATGAGTGGCGCTTGTCTTGGCTCTCTC 2132
QY 848 CAACGTGCTCTTCCACGCGCGCGCGCTTACGAGGCGCTTACGAGGCGCTTACGAGGCG 907
DB 2133 CCAACAGCGGCTTCTTCAAGGCTTGGGCGGCGGAGCTTACGAGGCGCTTACGAGGCG 2192
QY 908 AGCGTTCGCGCTTCTTGGGAGTCTTGGCGCTTCTTACGAGGCGCTTACGAGGCGCTTAC 967
DB 2193 CTTCTGCTGCTTACGAGGAGGATGCGCGGCAAGCTATGAGTGGCGCTTGTCTTGGCTCT 2252
QY 968 GGTCCGCTTACGAGGCTTCTTGGGAGTCTTGGCGCTTCTTACGAGGCGCTTCTTGGCT 1027
DB 2253 TCTGCGCGGAGTCTTCTTGGGAGTCTTGGCGGAGGATGAGCATGAGTGGCGGCGCTTGT 2312
QY 1028 GGCACCGGCGCTTCTTGGGAGTCTTGGCGGAGGATGAGCATGAGTGGCGGCGCTTGT 1052
DB 2313 GACCAACCGGATCTTACCGGATCTTCTTGGGAGTCTTGGCGGAGGATGAGCATGAGT 2337

RESULT 13

US-11-087-100-37/c
Sequence 37, Application US/11087100
Publication No. US20050266440A1

GENERAL INFORMATION:
APPLICANT: Mett, James
APPLICANT: Barclay, William
APPLICANT: Platt, James
APPLICANT: Kumer, Jerry
TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORPA of a PUFA Polyketide Synthase
FILE REFERENCE: 2997-29
CURRENT APPLICATION NUMBER: US/11/087,100
CURRENT FILING DATE: 2005-03-21
PRIOR APPLICATION NUMBER: 09/231,899
PRIOR FILING DATE: 1999-01-14
PRIOR APPLICATION NUMBER: 60/284,066
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 60/298,796
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/323,269
PRIOR FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.3
SEQ ID NO 37
LENGTH: 3886
TYPE: DNA
ORGANISM: Schizochytrium sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (2115)..(2115)
OTHER INFORMATION: n = a, c, g, or t

US-11-087-100-37

Query Match 2.9%; Score 41; DB 8; Length 3886;
Best Local Similarity 52.0%; Pred. No. 1;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 702 CTGAGAGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
DB 3850 CTGAGAGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3791
QY 762 CTTGGCGGAGGCTTCTTGGGAGTCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 821

DB 3790 ACTAGTATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3731
QY 822 GTGGGCTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
DB 3730 ATTGGGCTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3674

RESULT 14

US-11-087-084-37/c
Sequence 37, Application US/11087084
Publication No. US20050273883A1

GENERAL INFORMATION:
APPLICANT: Mett, James
APPLICANT: Barclay, William
APPLICANT: Platt, James
APPLICANT: Kumer, Jerry
TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORPA of a PUFA Polyketide Synthase
FILE REFERENCE: 2997-29
CURRENT APPLICATION NUMBER: US/11/087,084
CURRENT FILING DATE: 2005-03-21
PRIOR APPLICATION NUMBER: 09/231,899
PRIOR FILING DATE: 1999-01-14
PRIOR APPLICATION NUMBER: 60/284,066
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 60/298,796
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/323,269
PRIOR FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.3
SEQ ID NO 37
LENGTH: 3886
TYPE: DNA
ORGANISM: Schizochytrium sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (2115)..(2115)
OTHER INFORMATION: n = a, c, g, or t

Query Match 2.9%; Score 41; DB 8; Length 3886;
Best Local Similarity 52.0%; Pred. No. 1;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 702 CTGAGAGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
DB 3850 CTGAGAGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3791
QY 762 CTTGGCGGAGGCTTCTTGGGAGTCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 821
DB 3790 ACTAGTATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3731
QY 822 GTGGGCTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
DB 3730 ATTGGGCTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3674

RESULT 15

US-11-087-085-37/c
Sequence 37, Application US/11087085
Publication No. US20050273884A1

GENERAL INFORMATION:
APPLICANT: Mett, James
APPLICANT: Barclay, William
APPLICANT: Platt, James
APPLICANT: Kumer, Jerry
TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORPA of a PUFA Polyketide Synthase
FILE REFERENCE: 2997-29
CURRENT APPLICATION NUMBER: US/11/087,085
CURRENT FILING DATE: 2005-03-21
PRIOR APPLICATION NUMBER: 09/231,899

